NORTH ATLANTIC TREATY ORGANIZATION

SCIENCE AND TECHNOLOGY ORGANIZATION

AC/323(HFM-269)TP/936

www.sto.nato.int

STO TECHNICAL REPORT **TR-HFM-269**

Combat Integration: Implications for Physical Employment Standards

(Intégration des femmes au combat : implications pour les normes physiques d'emploi)

Final report of HFM-269.

Published December 2019

Distribution and Availability on Back Cover

NORTH ATLANTIC TREATY ORGANIZATION

SCIENCE AND TECHNOLOGY ORGANIZATION

AC/323(HFM-269)TP/936

www.sto.nato.int

STO TECHNICAL REPORT **TR-HFM-269**

Combat Integration: Implications for Physical Employment Standards

(Intégration des femmes au combat : implications pour les normes physiques d’emploi)

Final report of HFM-269

The NATO Science and Technology Organization

Science & Technology (S&T) in the NATO context is defined as the selective and rigorous generation and application of state-of-the-art, validated knowledge for defence and security purposes. S&T activities embrace scientific research, technology development, transition, application and field-testing, experimentation and a range of related scientific activities that include systems engineering, operational research and analysis, synthesis, integration and validation of knowledge derived through the scientific method.

In NATO, S&T is addressed using different business models, namely a collaborative business model where NATO provides a forum where NATO Nations and partner Nations elect to use their national resources to define, conduct and promote cooperative research and information exchange, and secondly an in-house delivery business model where S&T activities are conducted in a NATO dedicated executive body, having its own personnel, capabilities and infrastructure.

The mission of the NATO Science & Technology Organization (STO) is to help position the Nations' and NATO's S&T investments as a strategic enabler of the knowledge and technology advantage for the defence and security posture of NATO Nations and partner Nations, by conducting and promoting S&T activities that augment and leverage the capabilities and programmes of the Alliance, of the NATO Nations and the partner Nations, in support of NATO's objectives, and contributing to NATO's ability to enable and influence security and defence related capability development and threat mitigation in NATO Nations and partner Nations, in accordance with NATO policies.

The total spectrum of this collaborative effort is addressed by six Technical Panels who manage a wide range of scientific research activities, a Group specialising in modelling and simulation, plus a Committee dedicated to supporting the information management needs of the organization.

- AVT Applied Vehicle Technology Panel
- HFM Human Factors and Medicine Panel
- IST Information Systems Technology Panel
- NMSG NATO Modelling and Simulation Group
- SAS System Analysis and Studies Panel
- SCI Systems Concepts and Integration Panel
- SET Sensors and Electronics Technology Panel

These Panels and Group are the power-house of the collaborative model and are made up of national representatives as well as recognised world-class scientists, engineers and information specialists. In addition to providing critical technical oversight, they also provide a communication link to military users and other NATO bodies.

The scientific and technological work is carried out by Technical Teams, created under one or more of these eight bodies, for specific research activities which have a defined duration. These research activities can take a variety of forms, including Task Groups, Workshops, Symposia, Specialists' Meetings, Lecture Series and Technical Courses.

The content of this publication has been reproduced directly from material supplied by STO or the authors.

Published December 2019

ISBN 978-92-837-2261-8

Table of Contents

List of Figures

List of Tables

List of Acronyms

1RM	1 Repetition Max
ABC	Activities-specific Balance Confidence
ABCP	Army Body Composition Program
AC	All Corps
ACFT	Army Combat Fitness Test (U.S. Army)
ACOG	American College of Obstetrics and Gynaecology
ACS	All Corps Soldier
ADA	Americans with Disabilities Act
ADEA	Age, Discrimination in Employment Act
AFT	Annual Fitness Test
AI	Adverse Impact
AMAP	As Many As Possible
AML	Additional Maternity Leave
APFT	Army Physical Fitness Test (U.S. Army)
AUS	Australia
AVU-IGF	General Application Examination – Individual Basic Skills
BAM	Military mountaineering certificate
BFOQ	Bona Fide Occupational Qualification
BFOR	Bona Fide Occupational Requirement
BFT	Basic Fitness Test
BL	Box Lift
BM	Battlefield Maneuver
BMFT	Basic-Military-Fitness-Tool
BMI	Body Mass Index
BMOQ	Basic Military Officer Qualification
BMQ	Basic Military Qualification
BPFT	Basic Physical Fitness Test
BSM	Military skier's certificate
CA	Canadian Army
CAF	Canadian Armed Forces
CAN	Canada
CASEVAC	Casualty Evacuation
CAT	Category
CCPG	General Physical Fitness (French)
CCPM	Military combat fitness (French)
CCPS	Specific Physical Fitness (French)
CFLRS	Canadian Forces Leadership and Recruit School
CFT	Combat Fitness Test
CHAMP	Comprehensive High-Level Activity Mobility Predictor
CI	Confidence Interval
CMG	General Muscle Fitness (France)
CMTFE	Common military task fitness evaluation
CMTS	criterion measure task simulations
CRA	Civil Rights Act
CRF	Cardiorespiratory fitness
CST	Common Soldier Tasks

CV	Cardiovascular
CVD	Cardiovascular disease
DAOD	Defense Administration Orders and Directives
D-LOC	Directed Level of Operational Capability).
DMST	Danish Military Speed Test
DND	Department of National Defense
DRA	Diastis Recti Abdomini
DSTO	Defence Science and Technology Organisation
DXA	Dual Energy X-Ray Absorptiometry
E2C	Escape to cover
ECR	Cardioresiratory Endurace (France)
EEOC	Equal Employment Opportunity Commission
EFL	Entry Fitness Level
EOD	Explosive Ordnance Disposal
EPCC	Electronic Performance Comparison Compendium
ERP	Event-related Brain Potent
F:M	Female:Male
FFO	Full Fighting Order
FOMO	Fear Of Missing Out
FORCE	Fitness for Operational Requirement of Canadian Forces Employment
GCC	Ground Close Combat
GCC	Ground Close Combat
GOQ	Genuine Occupational Qualification
GXT	Graded Exercise Treadmill Test
HD	Higher Demands
HERO	Health Enhancement Research Organization
HRpQCT	High Resolution Peripheral Quantitative Computed Tomography
IBTS	Individual Battle Task Standards
ICC	intra-class correlation coefficient
ICD	International Classification of Diseases
IET	Initial Entry Training
IMTP	Isometric Mid-Thigh Pull
IST	Initial Strength Test
JTA	Job Task Analysis
kg	Kilogram
km	Kilometer
L&C	Lift and Carry
LCFT	Land Combat Fitness Test
MAPS	minimum acceptable performance standard
MBR	Military Basic Requirements
MBT	Medicine Ball Throw
MCS	Military Occupational Specialty Classifications Standards
MDC	Minimal Detectable Change
MEB	Medical Evaluation Board
MEL	Medical Employment Limitation

METS	Metabolic Equivalents
min	Minute
MJP	Military Judgement Panel
mL	Millilitre
MLC	Maximum Lifting Capacity
MMH	Manual Material Handling
MOBP	Method of Best Practice
MOET	Maritime Operational Evaluation Team
MOS	Military occupational specialty
MP	Military Police
MSFT	Multi-stage Fitness Test
MSKI	Musculoskeletal Injury
MSPS	Military Occupational Specific Physical Standards
MVC	Maximum Voluntary Contraction
MWL	Maximum Weight Limit
NAF	Netherlands Armed Forces
NBC	Nuclear Biological Chemical
NCO	Non-Commissioned Officers
NHRC	Naval Health Research Center
OFCCP	Office of Federal Contract Compliance Programs
OFT	Operational fitness test
OLP	Ordinary Least Product
OLS	Ordinary Lease Squared
OPAT	Occupational Physical Assessment Test
OPAT	Occupational Physical Assessment Test
OPR	Operational Physical Requirements
OSOR	Occupationally-Specific, Operationally-Relevant
P3T	Pregnancy and Post-partum Physical Training (P3T) program
PAR-Q	Participant Activity Readiness Questionnaire
PCA	Principle Component Analysis
PEB	Physical Evaluation Board
PES	Physical Employment Standard
PESA	Physical Employment Standard Army
PFT	Physical Fitness Test
PMT	Project Management Team
POI	Programs of Instruction
PSS(R)	Physical Selection Standards (Recruits)
PTI	Physical Training Instructor
RAFFT	Royal Air Force Fitness Tests
RMT	Representative Military Tasks
RN	Royal Navy
RNFT	Royal Navy Fitness Test
RNZAF	Royal New Zealand Air Force
RNZN	Royal New Zealand Navy
RR	Relative Risk
SAS	System Analysis and Studies Panel
SAT-PRO	Satisfaction with Prosthesis
SD	Standard Deviation
SEM	Standard Error of the Mean

SEPTA	Southeastern Pennsylvania Transportation Authority
SME	Subject Matter Expert
SNCO	Staff Non-Commissioned Officers
SOF	Special Operations Forces
SSSSW	Single Service Standard Setting Workshops
SW	Special Warfare
T&R	Training and Readiness
TAPES-R	Trinity Amputation and Prosthesis Experience Scales – Revised
TCA	Task Criticality Analyses
TECOM	Training and Education Command
TPC	Test de Puissance du Combatant
TRIAGE	Technique for Research of Information by Animation of a Group of Experts
TSSSW	Tri-Service Standard Setting Workshops
US	United States
UK	United Kingdom
USA	United States of America
USAF	United States Air Force
VDU	Visual Display Unit
VO2max	Maximum Amount of Oxygen the Body Can Utilize in a Specified Period
WASL	Working Above Shoulder Level
WHO	World Health Organization
WTBD	Warrior Tasks and Battle Drills

HFM-269 Membership List

CHAIR

Dr. Tara REILLY
Department of National Defence
CANADA
Email: tara.reilly@forces.gc.ca

MEMBERS

Dr. Neal BAUMGARTNER
US Air Force
UNITED STATES
Email: neal.baumgartner@us.af.mil

Dr. Sam BLACKER
University of Chichester
UNITED KINGDOM
Email: s.blacker@chi.ac.uk

Dr. Pieter BROWN
UK Royal Navy
UNITED KINGDOM
Email: NAVYINM-EMSAP2@mod.uk

Dr. Jace DRAIN
Defence Science and Technology Group
AUSTRALIA
Email: jace.drain@dsto.defence.gov.au

Lt Col E. Anne FIELDHOUSE
UK Ministry of Defence
UNITED KINGDOM
Email: anne.fieldhouse373@mod.gov.uk

Dr. Stephen FOULIS
US Army Research Institute of Environmental Medicine
UNITED STATES
Email: stephen.a.foulis.civ@mail.mil

Dr. Karl FRIEDL Col (ret.)
US Army
UNITED STATES
Email: friedlke@gmail.com

Prof. Julie GREEVES
Army Personnel Research Capability Organization
UNITED KINGDOM
Email: julie.greeves143@mod.uk

Dr. Henriette HASSELSTROM
Danish Armed Forces Health Services
DENMARK
Email: FSK-CM-MKS04@mil.dk

Keith HAURET
Army Public Health Center
UNITED STATES
Email: keith.g.hauret.civ@mail.mil

Dr. Pieter HELMHOUT
Netherlands Armed Forces
NETHERLANDS
Email: ph.helmhout.01@mindef.nl

Dr. Rachel IZARD
British Army
UNITED KINGDOM
Email: Rachel.izard715@mod.uk

Lt Col Sarah JACKSON (ret)
Army Pers Strat
UNITED KINGDOM
Email: drsarahjackson@live.co.uk

Mrs. Helen KILDING
New Zealand Defence Force
NEW ZEALAND
Email: helenk@dta.mil.nz

Prof. Dr. Col Dieter LEYK
Ministry of Defense
GERMANY
Email: dieterleyk@bundeswehr.org

Dr. Alexandra MALGOYRE Lt Col
Insitut de Recherche Biomédicale des Armées
FRANCE
Email: a.malgoyre@yahoo.fr

Prof. Stephen MYERS
University of Chichester
UNITED KINGDOM
Email: s.myers@chi.ac.uk

Dr. Tom O'LEARY
UK Ministry of Defence
UNITED KINGDOM
Email: Thomas.OLeary100@mod.gov.uk

Marilyn SHARP
US Army Research Institute of Environmental Medicine
UNITED STATES
Email: marilyn.a.sharp.civ@mail.mil

Hilde Kristin TEIEN
Forsvarets Forskningsinstitutt
NORWAY
Email: hilde-kristin.teien@ffi.no

Sqn Ldr James TREWEEK
Royal Air Force
UNITED KINGDOM
Email: james.treweek134@mod.gov.uk

ADDITIONAL CONTRIBUTORS

Dr. Anders AANDSTAD
The Norwegian Defence University College
NORWAY
Email: anders.aandstad@nih.no

Mr. Brian MCGUIRE
United States Marine Corps
UNITED STATES
Email : Brian.j.mcguire@usmc.mil

Maria C. CANINO
US Army Research Institute of Environmental Medicine
UNITED STATES
Email: maria.c.canino.ctr@mail.mil

Michael CAO
Department of National Defence
CANADA
Email: Michael.cao@forces.gc.ca

Dr. Deborah L. GEBHARDT
Human PRO
UNITED STATES
Email: dgebhardt@humrro.org

Dr. Hans Christian TINGELSTAD
Department of National Defence
CANADA
Email: Hanschristian.tingelstad@forces.gc.ca

Dr. Alex RAWCLIFFE
UK Ministry of Defence
UNITED KINGDOM
Alex.Rawcliffe103@mod.gov.uk

Dr. Olav Vikmoen
Forsvarets Forskningsinstitutt
NORWAY
Email: olav.vikmoen@ffi.no

PANEL/GROUP MENTOR

Ms. Eugenia KALANTZIS
Department of National Defence
CANADA
Email: Eugenia.kalantzis@forces.gc.ca

Combat Integration: Implications for Physical Employment Standards

(STO-TR-HFM-269)

Executive Summary

The purpose of this report is to publish the results of the activity by the NATO RTG HFM-269 panel between 2016 and 2019, on combat integration and the implications for Physical Employment Standards (PES). At the time the NATO RTG HFM-269 panel formed, all ten nations that were represented had either already lifted, or were planning to lift, exclusions of women joining combat roles within their Armed Forces. A fundamental element that has supported the opening of combat roles to women has been the development of role-related, age- and sex-free PES. Physical Employment Standards are physical fitness tests that are derived from the job-role that an individual performs. As such, when the physical requirements of a job-role are the same for all personnel, then the PES related to these roles should be the same irrespective of a person's sex, race or age.

- Chapter 1 is an introduction and summarises the key topics and interlinking themes in the report.
- Chapter 2 provides an overview of the current PES practices used for the selection and retention of military personnel in various international Armed Forces across the NATO nations represented on the NATO RTG HFM-269 panel.
- Chapter 3 presents a NATO guide for PES development supported by case studies of PES developments, example legal cases relating to PES and a series of infographics.
- Chapter 4 addresses potential incentivization methods that military services may use to motivate military service personnel to meet and exceed physical employment standards and enhance the quantity and quality of physical training.
- Chapter 5 considers the biological limitations to task performance and trainability in relation to PES and combat integration.
- Chapter 6 describes the role of the PES in musculoskeletal injury prevention.
- Chapter 7 discusses the implications for PES on sex differences in the physiological responses to prolonged military work.
- Chapter 8 summarises information on other topics relating to PES. Return-to-duty physical recovery timelines and policies are described with special considerations for post-partum women, post-injury recovery, and provisions for amputees following rehabilitation. In addition, the concept of the "soldier first" principle is reviewed.

Future efforts include consideration of how the model developed for PES might be useful in establishing future cognitive employment standards, and the role of physiological monitoring technologies (RTG HFM-260, Wearable physiological monitoring).

At the summation of this 3-year RTG, the unanswered questions appeared to lie in four genres:

1) The feasibility and fidelity of using PES as a predictor of musculoskeletal injury risk for military occupations;

2) Further consideration of women in the workplace and specific evidence-based recommendations for PES and physical training guidelines including pre- and post-partum;

3) The physical and occupational requirements of cyber operators; and

4) PES that factors in the effects of typical environmental and occupational exposures on military performance, including fatigue, sleep restriction, prolonged work, and thermal and hypoxic extremes.

The output from the NATO RTG HFM-269 panel will enhance NATO's military readiness by improving the understanding of how PES can be an important tool to support combat integration and more broadly improve health, reduce injury and optimise physical performance of both men and women across a diverse range of military settings. The future considerations for PES and integration with the outputs from other RTGs demonstrate that PES a continually evolving integrated process to support the optimisation of military readiness.

Intégration des femmes au combat : implications pour les normes physiques d'emploi

(STO-TR-HFM-269)

Synthèse

L'objet du présent rapport est de publier les résultats de l'activité du RTG HFM-269 de la Commission HFM de l'OTAN entre 2016 et 2019, portant sur l'intégration des femmes au combat et ses implications pour les normes physiques d'emploi (PES). Au moment de la formation du RTG HFM-269, les dix pays représentés avaient déjà levé ou prévoyaient de lever l'interdiction faite aux femmes de prendre part au combat dans leurs forces armées. L'ouverture des rôles de combattant aux femmes a été soutenue par un élément fondamental, le développement de PES relatives aux rôles, non sexuées et sans limite d'âge. Les normes physiques d'emploi sont des tests d'aptitude physique qui découlent du rôle professionnel d'un individu. À ce titre, lorsque les exigences physiques d'un rôle professionnel sont les mêmes pour tout le personnel, les PES relatives à ces rôles sont indépendantes du sexe, de l'origine ethnique ou de l'âge.

- Le chapitre 1 est une introduction et résume les sujets essentiels et les thèmes connexes du rapport.
- Le chapitre 2 donne une vue d'ensemble des pratiques actuelles liées aux PES pour sélectionner et conserver le personnel militaire dans diverses forces armées internationales des pays de l'OTAN représentés dans le RTG-269 HFM.
- Le chapitre 3 présente un guide OTAN d'élaboration des PES, étayé par des études de cas sur l'élaboration de PES, des exemples d'affaires juridiques liées aux PES et une série d'infographies.
- Le chapitre 4 traite des méthodes potentielles que les services militaires peuvent utiliser pour motiver le personnel militaire en service à atteindre et dépasser les normes physiques d'emploi et améliorer la quantité et la qualité de l'entraînement physique.
- Le chapitre 5 étudie les limites biologiques à l'exécution des tâches et à l'aptitude à l'entraînement, en rapport avec les PES et l'intégration des femmes au combat.
- Le chapitre 6 décrit le rôle des PES dans la prévention des traumatismes musculo-squelettiques.
- Le chapitre 7 discute des implications, pour les PES, des réponses physiologiques différentes des deux sexes face au travail militaire prolongé.
- Le chapitre 8 résume les informations se rapportant à d'autres sujets relatifs aux PES. Nous décrivons les délais et les politiques de rétablissement physique avant la reprise du service, en accordant une considération particulière aux femmes post-partum, au rétablissement après une blessure et aux amputés après la réadaptation. Nous étudions de plus le concept de « l'universalité du service ».

Les travaux futurs devraient porter sur la manière dont le modèle développé pour les PES pourrait être utile à l'établissement de futures normes cognitives d'emploi et sur le rôle des technologies de monitorage physiologique (RTG HFM-260, Monitorage physiologique portable).

En conclusion de ce RTG de trois ans, les questions sans réponse sont de quatre ordres :

1) La faisabilité et la fiabilité d'utilisation des PES pour prédire le risque de traumatisme musculo-squelettique d'activités militaires ;

2) La prise en considération plus poussée des femmes sur le lieu de travail, les recommandations spécifiques pour les PES, basées sur des éléments tangibles, et les principes directeurs d'entraînement physique incluant la période et pré et post-partum ;

3) Les exigences physiques et professionnelles applicables aux cyberopérateurs ; et

4) Les PES qui tiennent compte des effets de l'environnement typique et des expositions professionnelles sur les performances militaires, y compris la fatigue, la restriction de sommeil, le travail prolongé et les extrêmes thermiques et hypoxiques.

Les résultats du RTG HFM-269 serviront l'état de préparation militaire de l'OTAN en améliorant la compréhension du rôle important que peuvent jouer les PES pour appuyer l'intégration des femmes au combat et plus largement, améliorer la santé, réduire les blessures et optimiser les performances physiques des hommes et des femmes dans une large palette de contextes militaires. Les futures considérations concernant les PES et l'intégration avec les résultats d'autres RTG démontrent que les PES sont un processus intégré en perpétuelle évolution qui soutient l'optimisation de l'état de préparation militaire.

Chapter 1 – INTRODUCTION

S. Blacker
University of Chichester
UNITED KINGDOM

K. Friedl
US Army Research Institute for Environmental Medicine
UNITED STATES

T.J. Reilly
Department of National Defence
CANADA

The purpose of this report is to publish the results of the activity by the NATO HFM RTG-269 panel between 2016 and 2019, on combat integration and the implications for Physical Employment Standards (PES). At the time the NATO RTG HFM-269 panel formed, all ten nations that were represented had either already lifted, or were planning to lift, exclusions of women joining combat roles within their Armed Forces. A fundamental element that has supported the opening of combat roles to women has been the development of role-related, age-, sex- free PES. Physical Employment Standards are physical fitness tests that are derived from the job-role that an individual performs. As such, when the physical requirements of a job-role are the same for all personnel, then the PES related to these roles will be the same irrespective of a person's sex, race or age.

Chapter 2 provides an overview of the current PES practices used for the selection and retention of military personnel in various international Armed Forces across the NATO nations represented on the NATO RTG HFM-269 panel. A number of commonalties are apparent between the PES across the different nations:

1) The PES are typically underpinned by a Job Task Analysis (JTA) to define the most physically demanding elements of job-roles.
2) The components of fitness required to safely and effectively perform the job-tasks are evaluated within the PES and cover Endurance (aerobic endurance, anaerobic capacity) Strength (muscular strength, muscular endurance, muscular power) and Mobility (flexibility, balance, speed, agility, coordination).
3) The tests within a PES are a combination of either one-person task simulations (e.g., loaded march, casualty drag, equipment carry) and / or simple gym-based tests (e.g., run, push-ups, medicine ball throw). The choice of the type of test(s) used varies between nations and is influenced by the requirements of military organisations, resources and the population to which they are being applied.
4) Most Armed Forces branches have developed role-related, sex-, and age-free PES underpinned by the principle of generic military tasks that all personnel within a role would be expected to complete (e.g., load carriage) and then included additional tests where specialist tasks exceed the physical demands of the generic military tasks.
5) The final standards or cut-scores associated with each of the PES are typically linked to the role and are therefore sex- and age-free.

Chapter 3 presents a NATO guide for PES development which is supported by case studies of PES developments, example legal cases relating to PES and a series of infographics. The guide first describes the need and process of establishing a project management team, military judgement panels and processes for PES governance. Five key steps in PES development are then described:

1) Job Task Analysis (JTA) – The JTA is the foundation on which a PES is developed and involves identifying, documenting and down selecting the essential/critical physically demanding tasks using a combination of subjective and objective methods.

2) Scenario Construction – Scenarios of the down selected essential/critical physically demanding tasks are constructed so that ergonomic analysis, physical and/or physiological measures can be made to concentrate, combine and eliminate tasks to inform the development of the tests within a PES.
3) Test Development – Tests are developed to simulate or measure the components of fitness and/or physical actions required to perform the job-tasks. Considerations are presented regarding balancing fidelity and feasibility, and analysis for bias and measures of maximum performance to establish correlations.
4) Setting Standards – A range of performance scores are generated during the test development and cut-scores for the tests within a PES are set. Points are discussed regarding adverse impact, predictive bias and accommodation.
5) Validation and Reliability – The processes establishing the validity and reliability of both predictive tests and task simulations are discussed.

Chapter 4 addresses potential incentivization methods that military services may use to motivate military service personnel to meet and exceed physical employment standards and enhance the quantity and quality of physical training. Section one provides a brief summary of the extensive literature on incentives used in the workplace. Sections two and three address sex-fair and sex-neutral tests to which military services may apply incentives. Finally section four addresses military-specific limitations and opportunities for incentivizing members to meet or exceed physical fitness standards and includes a summary table of incentive methodology.

Chapter 5 considers the biological limitations to task performance and trainability in relation to PES and combat integration. Physical and physiological sex differences pose one of the greatest barriers to incorporating large numbers of women into the combat arms professions. Data is summarised in Chapter 4 that show that the average man tends to outperform the average woman on PES tests but there is always a degree of overlap in that some women will outperform some men. Evidence is presented showing that some physical fitness tests and task-based tests are more biased than others and that sex differences can vary greatly across tests of the same physical capacity. Finally, physical training strategies are discussed that may reduce sex differences on PES.

Chapter 6 describes the role of the PES in musculoskeletal injury prevention. Data are presented on the incidence and risk factors for musculoskeletal injury in the military in relation to sex and in physical fitness. The current role of PES in musculoskeletal injury prevention and future directions of research across a selection of the nations represented on the NATO RTG HFM-269 panel are discussed. The chapter concludes that the introduction of PES which reflects the physical demands of a job-role should result in a reduction in MSKI risk as a person-job fit is achieved at selection and maintained throughout a career. This will impact on what physical training service personnel undertake and should therefore be protective. As NATO militaries develop and implement PES, it is imperative that they develop an injury surveillance program, establish baseline injury rates, and then carefully monitor trends in injury rates, types, causes, and outcomes such as restricted duty after implementation of the PES.

Chapter 7 discusses the implications for PES on sex differences in the physiological responses to prolonged military work. PES are typically performed in a controlled environment. However, the job-tasks from which PES attempt to replicate are often performed during prolonged military work in a multi-stressor environment may include inadequate sleep, high physical activity levels, inadequate energy intake, psychological strain, and a wide range of environmental conditions. A PES primarily evaluates acute (< 4 hours) physiological performance and does not directly test the ability to withstand all the military stressors over repeated days.

There is limited research available comparing sex differences in relation to these topics. Based on this limited evidence, Chapter 7 concludes that compared to men, women appear to experience greater physiological strain for the same amount of physical activity in military settings, likely due to on average lower muscle mass and aerobic fitness. However, it appears that women experience smaller reductions in absolute lean body mass, lower muscle fatigue and improved recovery compared to men following prolonged military work. Further research is recommended to examine sex differences in physical and physiological performance following exposure to the same periods of prolonged military work.

Chapter 8 summarises information on other topics relating to PES. Return-to-duty physical recovery timelines and policies are described with special considerations for post-partum women, post-injury recovery, and provisions for amputees following rehabilitation. In addition, the concept of the "soldier first" principle is reviewed. Examples are presented and considerations discussed of how selected nations on the NATO RTG HFM-269 panel are currently or planning to use PES in relation to these topic areas.

Future efforts building on this work were also considered by this panel. A two-day symposium following the final panel meeting at the US Military Academy at West Point, New York, considered how the model developed for PES might be useful in establishing future cognitive employment standards. Rather than identifying any specific occupationally-relevant domains of brain and behavior function, the ability to function as part of a team emerged as a primary selection criterion for military task performance. Common misperceptions about cyber warrior physical demands were corrected with clarification of the role of the Signals officers in combat roles.

The role of physiological monitoring technologies which are becoming ubiquitous in modern society was also considered, with information presented from another HFM panel (RTG HFM-260, Wearable Physiological Monitoring). Two key applications can be anticipated: the longitudinal assessment of physical demands to determine different demands with changes in roles; and for the improvement of work safety and efficiency, with the acquisition of actual data on frequency and intensity of specific physical demands. Longitudinal assessments come with cautions about intruding on personal health information unrelated to occupational capacity and inappropriate surveillance of worker behavior.

At the summation of this 3-year RTG, the unanswered questions appeared to lie in four genres:

1) The feasibility and fidelity of using PES as an indicator of MSKI risk for military occupations.
2) The specific health and fitness considerations of female forces, particularly the need for evidence-based recommendations for PES and physical training guidelines including pre- and post-partum.
3) The physical and occupational requirements of future cyber operators to maximise their effectiveness and resilience in the workplace.
4) How to consider the consequences of fatigue (sleep deprivation and long field exposure) as well as hot climate on PES to potentially adjust the level required depending on the place where the task is performed.

These four topics were relatively new to the ten nations represented here as there was no consensus on the specifics of how to best collect and perform longitudinal epidemiological research to identify the utility of PES in injury prevention, best practice for return to work postpartum, or the development of a PES for cyber operators, or environmental exposure.

Complementary NATO HFM included SAS 137 Integration of Women into Ground Combat Units, now closed, which dealt with the social and cultural considerations of Combat Integration. NATO RTG HFM-283 Reducing Musculo-Skeletal Injuries (2017 – 2020) is focused on primary preventive measures to reduce MSKI by:

1) Promoting the sharing of information among participating nations;
2) Identifying the causes and associated risk factors for MSKI;

3) Identifying existing and novel strategies/technologies which may reduce the injury burden; and
4) Linking to other on-going STO-activities.

The interaction of PES with MSKI is not covered in this group, however USA, Canada, UK, and Australia are performing longitudinal studies on this topic.

The information in the present report provides guidance and examples for establishing, implementing, using and refining PES. The output from the NATO RTG HFM-269 panel will enhance NATO's military readiness by improving the understanding of how PES can be an important tool to support combat integration and more broadly improve health, reduce injury and optimise physical performance of both men and women across a diverse range of military settings. The future considerations for PES and integration with the outputs from other RTGs demonstrate that PES is a continually evolving integrated process to support the optimisation of military readiness.

Chapter 2 – REVIEW OF CURRENT INTERNATIONAL PES PRACTICES

S.D. Blacker
University of Chichester
UNITED KINGDOM

H. Kilding
New Zealand Defence Force
NEW ZEALAND

T.J. Reilly
Department of National Defence
CANADA

M. Sharp, and S. Foulis
US Army Research Institute for Environmental Medicine
UNITED STATES

B.J. McGuire
US Marine Corps
UNITED STATES

P. Brown
Institute of Naval Medicine
UNITED KINGDOM

J. Greeves
Army Personnel Research Capability Organization
UNITED KINDOM

A. Fieldhouse
UK Ministry of Defence
UNITED KINGDOM

S. Myers
University of Chichester
UNITED KINGDOM

J. Treweek
Royal Air Force (RAF)
UNITED KINGDOM

J. Drain
Defence Science and Technology Group
AUSTRALIA

D. Leyk
Central Institute of the Bundeswehr Medical Service
GERMANY

A. Malgoyre
Institut de Recherche Biomédicale des Armées
FRANCE

H. Hasselstrøm
Danish Defence Medical Command
DENMARK

H. Teien
Norwegian Defence Research Establishment (FFI)
NORWAY

A. Aanstad
The Norwegian Defence University College
NORWAY

P. Helmhout
Royal Netherlands Army
NETHERLANDS

2.1 INTRODUCTION

This chapter provides an overview of the current Physical Employment Standard (PES) practices used for the selection and retention of military personnel in various international Armed Forces branches across the NATO nations that have contributed to this technical report. The focus has been placed on branches of the Armed Forces in the nations that currently use, or are developing, role-related PES (i.e., tests that are not reliant on normative or health-related standards). Where possible, a summary of the research process and/or evidence base which underpins the PES and the application of the tests and standards have been provided. The information presented in this chapter is intended to provide examples of how the methods for PES development described in Chapter 3 have been used and the resultant tests and standards that have been implemented. Table 2-1 provides a high-level summary of the points during personnel's career at which

International Armed Forces branches currently utilise Physical Employment Standards and Table 2-2 and Table 2-3 summarize the tests used at each stage. Subsequent sections of this chapter provide a detailed description of how the PES have been developed and are utilised by each group.

Table 2-1: Simple Summary of International Armed Forces Branches Use of Physical Employment Standards (PES) and Non-PES Relate Fitness Standards (Where PES are Not Used) Through Career.

Armed Forces Branch	Pre-Employment / Selection	In- Service / Annual
Canadian Armed Forces (Tri-Service)	PES	PES
US Army	PES	PES
US Marine Corp	PES	PES
UK Royal Navy	PES *and Non-PES Related Standards*	PES *and Non-PES Related Standards*
British Army	PES	PES
UK Armed Forces – Ground Close Combat	PES	PES
Royal Air Force	*Non-PES Related Standards*	*Non-PES Related Standards*
Australian Army	*Non-PES Related Standards*	PES
New Zealand Army	*Non-PES Related Standards*	PES and *Non-PES Related Standards*
Royal New Zealand Navy	*Non-PES Related Standards*	PES
Royal New Zealand Air Force	*Non-PES Related Standards*	PES
Bundeswehr (Germany)	PES	PES
French Armed Forces	*Non-PES Related Standards*	PES
Denmark (Tri-Service)	None / PES (role dependent)	PES
Norwegian Armed Forces (Tri-Service)	PES	PES
Netherlands (Tri-Service)	PES	*Non-PES Related Standards*

Table 2-2: Detailed Summary of PES International Armed Forces Use of Physical Employment Standards Used to Screen Applicants at the Point of Pre-Employment / Selection.

Armed Forces Branch	Details
Canadian Armed Forces (Tri-Service)	Fitness for Operational Requirements of the CAF Employment (FORCE) Evaluation – Four fitness tests (20 m rushes, sandbag lift, intermittent laded shuttles and sandbag drag) with a common tri-service standard. Designed to infer performance of the 6 essential job tasks within the Common Military Task Fitness Evaluation (CMTFE), based on analysis of physical demands.
US Army	Pass the Army Physical Fitness Test at 60% of the age- and gender-adjusted score. Army Combat Fitness Test (in development by Training and Doctrine Command); Wearing shorts and t-shirt (deadlift, standing power throw, hand-release-push-up, 25 m Sprint / drag / carry, hanging leg tuck, 2-mile run). Expected rollout in Fall 2020.

Armed Forces Branch	Details
US Army (cont'd)	Soldier Readiness Test (in development by Army Forces Command); Wearing body armour complete the following in under 23 min: 220 lb tire flips, agility test, 240 lb dummy drag, sand bag stack, sand bag toss, 1.5-mile run on unimproved terrain.
US Marine Corp	A combination of gender-free and gender-fair standards on four tests; pull-ups, 1.5-mile run, crunches and pull-ups.
UK Royal Navy	Applicants must achieve an age and gender-fair field-based aerobic fitness standard that is within 10% of the in-service PES standard (see Table 2-3 UK Royal Navy – Part 1). Research is currently being undertaken to develop a new pre-employment PES for Royal Navy seafarers, which is due to be completed by 2021.
British Army	Applicants complete three gym-based predictor tests; 4 kg medicine ball throw, Isometric Mid-Thigh Pull (IMTP) and a 2 km run. The pass standards on each test varies depending on the physical requirements of the military job role to which the individual is applying and account for changes in physical performance during training to the point where the individual is expected to achieve their role-related output standard.
UK Armed Forces Ground Close Combat (GCC) Roles	Applicants to the British Army and RAF GCC roles complete three gym-based predictor tests; 4 kg medicine ball throw, Isometric Mid-Thigh Pull (IMTP) and a 2 km run. The pass standards on each test varies depending on the physical requirements of the military job role to which the individual is applying and account for changes in physical performance during training to the point where the individual is expected to achieve their role-related output standard.
Royal Air Force	**Non-PES related**. Multi-Stage Fitness Test, 1-min press-ups, 1-min sit-ups. Tests are age- and gender-fair.
Australian Army	**Non-PES Related**. Complete a 20 m multi-stage shuttle run test (gender-free), push-ups (gender-fair), sit-ups (gender-free). Not linked to a JTA or In-Service tests.
New Zealand Army	**Non-PES Related**. Candidates must pass a Multi-Stage Fitness Test (MSFT), curl-ups and press-ups tests (all with gender-fair standards). Results are graded from "minimum pass" to "strong pass" and the higher the grading, the more competitive the candidate will be in the overall selection process. On Induction day, candidates must attain the NZ Army Entry Fitness Level (EFL) which is a 2.4 km run, curl-up and press-up test with gender-fair standards.
Royal New Zealand Navy	**Non-PES Related**. Candidates must pass a Multi-Stage Fitness Test (MSFT), press-up test (both with gender-fair standards) and an Entry Swim Test (gender-free). Results are graded from "minimum pass" to "strong pass" and the higher the grading, the more competitive the candidate will be in the overall selection process. On Induction day, candidates must repeat and pass the entry test.
Royal New Zealand Air Force	**Non-PES Related.** Candidates must pass a Multi-Stage Fitness Test (MSFT) and press-up test (both with gender-fair standards). Results are graded from "minimum pass" to "strong pass" and the higher the grading, the more competitive the candidate will be in the overall selection process.

Armed Forces Branch	Details
Bundeswehr (Germany)	Gender-free tests and minimum standards on three tests; 11 x 10 m shuttle run (≤ 60 sec), flexed arm hang in the chin up position (≥ 5 sec), bicycle ergometer test cycling at 130 W and 80 rpm for minimum 6:30 min. But a gender and age-related grading system.
French Armed Forces	**Non-PES Related**. Candidate must pass a shuttle test to assess aerobic capacity; a minimum of 3 pull-ups for men and high rowing bar for women and a coordination circuit.
Denmark (Tri-Service)	PES (CAT1) is used as admission requirements different educations in the Danish Armed Forces. There are higher demands for the SOF and MP – and the exercises in the test is somewhat different.
Norwegian Armed Forces (Tri-Service)	Gender-free tests and standards on four tests; maximal walk / run on treadmill, seated medicine ball put (10 kg), standing long jump, pull-ups (vertical or horizontal).
Netherlands (Tri-Service)	A series of role-related, gender-free tests; Loaded march (20 to 48 min with 25 to 45 kg, depending on role), repeated lifting (both floor level and height-adjusted table level) and carrying of an ammunition box 25 m at chest height (20 / 30 / 40 kg; weights and numbers of repetitions depending on job cluster). Digging sand from one compartment to the other in a bisectional sand bucket (1 m^3) for 1 to 2 min, depending on job cluster. 'Operational endurance' assessment with 12 min run test, with cluster-dependent standards (2200 to2700 m). Working above shoulder level (cluster independent): moving a weight (both hands) from one shaft to another for 1 min at a body height-adjusted height, followed by screwing and unscrewing of a wingnut (left and right hand) for 1 min at the same height. Indoor test series (cluster independent; clambering and scrambling on and off obstacles; squat for 5 seconds; shooting positions: lying, kneeling left / right; 6 m crawl on hands and knees; 6 m military crawl.

Table 2-3: Detailed Summary of PES International Armed Forces Use of Physical Employment Standards Used to Screen Applicants at the Point of Pre-Employment / Selection.

Armed Forces Branch	Details	Frequency
Canadian Armed Forces (Tri-Service)	Fitness for Operational Requirements of the CAF Employment (FORCE) Evaluation – Four fitness tests (20 m rushes, sandbag lift, intermittent laded shuttles and sandbag drag) with a common tri-service standard. Designed to infer performance of the six essential job tasks within the Common Military Task Fitness Evaluation (CMTFE), based on analysis of physical demands.	Yearly
US Army	Pass the Army Physical Fitness Test at 60% of the age and gender adjusted score. Army Combat Readiness Test (in development by Training and Doctrine Command); Wearing shorts and t-shirt (deadlift, standing power throw, T-push-up, 25m Sprint/drag/carry, hanging leg tuck, 2-mile run). Expected rollout in 2018 / 2019.* Soldier Readiness Test (in development by Army Forces Command); Wearing body armour complete the following in under 23 min: 220 lb tyre flips, agility test, 240 lb dummy drag, sand bag stack, sand bag toss, 1.5-mile run on unimproved terrain.	Yearly

Armed Forces Branch	Details	Frequency
US Marine Corp	All Marines must take the PFT between 1 Jan – 30 June, and the CFT 1 July – 31 Dec every calendar year.	Yearly
UK Royal Navy	The Royal Navy Fitness Test (RNFT). Part 1 – an age and gender-fair field-based aerobic fitness test, the origins of which (in-part) derive from the physical demands of Naval fire fighting tasks. Assessments include either the 20 m Multi-Stage Fitness Test, 2.4 km run, or the 1.6 km Rockport Walk Test. Part 2 – an age and gender-free test that simulates the resupply of fire fighting foam drums. Involves carrying × 2, 20 kg power bags a distance of 60 m. Research is currently being undertaken to develop a new in-service PES for Royal Navy seafarers, which is due to be completed by 2020.	Yearly
British Army	The British Army Non-GCC roles currently complete an 8-mile Loaded March carrying between 15 – 25 kg depending on the physical demands of their job role. Research is currently being undertaken to develop new PES for the British Army Non-GCC roles which is due to be completed in 2021.	Yearly
UK Armed Forces Ground Close Combat (GCC) Roles	UK Armed Forces GCC personnel must annually complete a series of Representative Military Tasks (RMTs) which are single person simulations of the job tasks that they perform in their roles. The RMTs and associated pass standards very depending on the job tasks and the physical requirements of each military job role in which the personnel serve.	Yearly
Royal Air Force	**Non-PES related**. In-Service test is a health-related test that measures aerobic capacity and muscular endurance. All personnel are required to undertake a Multi-Stage Fitness Test, 1-min press-up and 1-min sit-up test. Standards are age- and gender-fair.	Yearly
Australian Army	Physical Employment Standards Army (PESA) Four role-related tests derived from a JTA with gender-free standards depending on employment category; forced march, fire and movement, lift and carry, box lift and place.	Yearly
New Zealand Army	New Zealand Army Land Combat Fitness Test (LCFT). Four role-related tests derived from a JTA with gender-free standards depending on employment category; tactical movement, fire and manoeuvre, lift and place, lift and carry.	Yearly
Royal New Zealand Navy	Royal New Zealand Navy fitness test has two components. Component 1 – a health-related fitness component (the Multi-Stage Fitness Test). Component 2 – Job-related fitness component; a lift and carry and a dummy drag.	Yearly
Royal New Zealand Air Force	The Royal New Zealand Air Force (RNZAF) Operational Fitness Test (OFT). Part 1 – push-up test, with different standards based on age and gender. Part 2 – a timed 5 km march carrying an evenly distributed weight of 20 kg.	Yearly
Bundeswehr (Germany)	Fundamental Fitness: Gender-free tests and minimum standards on three tests; 11 x 10 m shuttle run (≤ 60 sec), flexed arm hang in the chin up position (≥ 5 sec), 1000m run (≤ 6:30 min). But a gender and age-related grading system.	Yearly

Armed Forces Branch	Details	Frequency
Bundeswehr (Germany) (cont'd)	Basic-Military-Fitness: Gender-free tests and standards; a loaded march in field uniform (no helmet, no ballistic body armour) with a 15 kg backpack. The mandatory level for all soldiers is to cover 6 km in 60 min. Additionally 100 m swimming fully dressed (swimming suit + field uniform pants + field uniform jacket) in 4 minutes with subsequent undressing.	Yearly
French Armed Forces	All Armed Forces Personnel (Tri-Service) -Complete 12 min cooper test or shuttle test or vameval in sport wear on track AND aquatic ease test AND sit-u, push-up and sit-up OR rope and sit-up and pull-up. All Army and Parachute personnel also complete an 8 km run with back pack in a mixed terrain (< 55 min). Mountain troop; Military skier's certificate (BSM) and Military mountaineering certificate or (BAM).	Yearly
Denmark (Tri-Service)	Basic Physical Fitness Test (BPFT) covers the Category 1 roles and has gender-free standards that have lower pass standards for older ages; 12 min run or shuttle run test, split squat, dips on a bench, horizontal pull-up and burpees. Personnel in more physically demanding roles (Categories 2 – 5) must complete an Operational Physical Requirements (OPR) based on their job role which are gender- and age-free. The OPR includes all the tests of the BPFT and the Danish Military Speed Test (DMST; intermittent shuttle running) and deadlift.	Yearly
Norwegian Armed Forces (Tri-Service)	All military personnel (irrespective of branch) complete four fitness tests with gender-free tests, but gender (and age) fair standards partly based on a JTA of the general fitness required by military personnel; 3000 m run (or 20 m shuttle run test), standing medicine ball put (10 kg); standing long jump; pull-ups (vertical or horizontal).	Yearly
Netherlands (Tri-Service)	**Non-PES Related**. The Netherlands Armed Forces (NAF) Physical Fitness Test, comprises of a 12-min run, push-ups, and sit-ups with gender- and age-dependent standards.	Yearly

The majority of International Armed Forces branches that have implemented or are developing role-related PES have broadly used the research process described in Chapter 3 supported by military judgement and/or Subject Matter Expert (SME) guidance. In summary the following commonalties are apparent:

1) The PES are typically underpinned by a Job Task Analysis (JTA) to define the most physically demanding elements of the military roles to which they will be applied.

2) Researchers have identified the components of fitness required to safely and effectively perform the job roles. Most military job roles require varying degrees of the components of fitness described in Figure 2-1; Endurance (aerobic endurance, anaerobic capacity) Strength (muscular strength, muscular endurance, muscular power) and Mobility (flexibility, balance, speed, agility, coordination). The precise definitions and groupings of these components of fitness vary between nations.

3) Tests within a PES are developed to measure these components of fitness in personnel through a combination of either one-person task simulations (e.g., loaded march, casualty drag, equipment carry) and/or simple gym-based tests (e.g., running, push-ups, medicine ball throw). The choice of the type of test(s) used varies between groups and is likely guided by the requirements of military organisations, resources and the population to which they are being applied. Task simulations have

more typically been adopted for trained personnel and where more resources are available, and/or in a more specialist populations (for example for Infantry roles compared to whole tri-service populations). Gym-based tests appear more typically used where large numbers of personnel are needed to be tested quickly and/or at the point of entry or selection where applicants don't possess the training and/or technical capabilities to safely complete the task simulations.

4) Most of the Armed Forces branches appear to have developed role-related, gender- and age- free PES underpinned by the principle of generic military tasks that all personnel within a role would be expected to complete (e.g., load carriage) and then included additional tests where specialist tasks exceed the physical demands of the generic military tasks.

5) The final standards or cut-scores associated with each of the PES are typically linked to the role and are therefore gender- and age-free. As a result, some PES have generic cut-scores across all roles and other PES have multiple tiers of cut-scores for the same task simulations or gym-based tests depending on the physical demanding tasks performed in a job role.

In several groups, gender and age fair PES have been retained, however these appear to be as a result of an amalgamation of health-related or incentivisation into the PES. Gender and age fair standards always result in lower requirements for women compared to men and for older compared to younger individuals. These standards are typically defined based on known relationships between physical fitness and health (e.g., the higher risk of cardiovascular disease associated with lower aerobic endurance) and/or the age-related declines in the components of fitness that are typically observed in normative population data. Gender and age fair standards go against the principles of role-related fitness where the job tasks remain the same irrespective of gender or age. No nations appear to be developing new generic gender and age fair norm fitness tests and are moving towards the development of role-related, gender-free PES. The following sections provide detailed descriptions of the PES adopted or being developed by each of the International Armed Forces branches listed in Table 2-1.

2.2 CANADIAN ARMED FORCES (TRI-SERVICE)

Armed Forces Branch	Canadian Armed Forces (CAF)
Approximate date current PES implemented	2014
Approximate duration of the PES	1 – 2 h
Test standards the same for gender	Yes
Test standards the same for all ages	Yes
Frequency of PES assessment	Point of selection / Annual;
PES user (e.g., tri-service, single-service or specialist)	Tri-service
Is the PES currently being reviewed or redeveloped?	No
If 'Yes', anticipated date when new PES be implemented	N/A

2.2.1 Summary of Current PES Processes

The Common Military Task Fitness Evaluation (CMTFE) reflects the Minimum Physical Fitness Standard for all Canadian Military Personnel across trade classification, age, and gender [1]. The evaluation consists of six essential, common, and physically demanding tasks based on extensive evidence-based research [2]. There are minimum performance standards on each of these tasks, required by all personnel as they are considered Essential, and serve as the basis in the establishment of Bona Fide Occupational Requirements (BFOR) in agreement with the Universality of Service Principle [3]. Table 2-4 highlights the six CMTFE (task simulations) along with respective standards.

NATO HFM RTG 269 Combat Integration: Implications for Physical Employment Standards

Fitness	Fitness Component	Definition	Example Activities
Endurance	Aerobic Capacity	Ability to sustain sub-maximal low-to-moderate/high intensity activity for a sustained period of time (minutes to hours), typically involving dynamic whole-body activities	Sustained patrolling carrying load (e.g. ≥ 30 kg) or digging a fire trench
	Anaerobic Capacity	Ability to sustain intermittent or continuous near maximal or maximal efforts for a short period of time (seconds to minutes), typically involving dynamic whole-body activities	Fire and movement task or a break contact task
Strength	Muscular Strength	Ability of a muscle group to exert maximal force in a single voluntary contraction (< 5 seconds)	Lifting objects, e.g. a casualty, equipment onto a vehicle. Standing up from kneeling while carrying a heavy load
	Muscular Endurance	Ability of a muscle group to repeatedly generate an intermittent or continuous moderate-to-high absolute force for a more prolonged period of time (seconds to minutes)	Repetitively lifting and carrying stores or a stretcher casualty evacuation
	Muscular Power	Ability to exert maximal external force in the shortest possible time (typically less than 1 second)	Breaking down a compound/ building door or jumping over a ditch or low wall
Mobility	Flexibility	The ability to voluntarily stretch, flex or lengthen parts of the body as far as possible i.e. the range of motion around a joint	Lifting a leg over a fence or bending down to pass under a low obstacle
	Balance	Maintenance of equilibrium while stationary or moving	Maintenance of a stable firing position
	Speed	Ability to perform movements in a short period of time	Rapid movement between fire positions
	Agility	Ability to change the position of the entire body in space with speed and accuracy	Hurdling a fence or rapidly changing running direction (e.g. fire and movement task)
	Coordination	Ability to synchronise the senses (e.g. sight/hearing) with body parts to move smoothly and accurately	Bringing weapons systems to bear and accurately engaging with the enemy

(Blacker S. et al, 2018)

NATO OTAN HUMAN PERFORMANCE HUMAINE R&D OPRG University of Chichester

Figure 2-1: The Components of Fitness that Underpin Military Job Task Performance.

Table 2-4: CMTFE Task and Standards.

Task	Standard
Escape to Cover – with rubber rifle, complete circuit involving (50m, 30m) sprints, stops, kneeling, and 10m of low crawl.	≤ 68 sec
Vehicle Extraction – manipulate 86 kg mannequin off of platform (76cm), drag (5m), then return 52 kg mannequin via carry back onto platform.	Pass / fail
Picking and Digging – Using sledgehammer to strike a beam to simulate picking action. Maintain 30 / 60 work to rest rate until beam is displaced total of 4m. After 2 min. rest, transfer 180kg of loose gravel into empty box by digging. Maintain 60 sec work to 30 sec rest ratio.	≤ 18 min of picking ≤ 18 min of digging
Stretcher Carry – Deadlift 43 kg, carry for 25m twice with 15 sec break. Separately, lift 21.5kg bar onto 91.5cm platform.	No stopping outside of rest period
Sandbag Fortification – Lift 60 sandbags (20kg) from pallet and place on top of 91.5cm platform.	≤ 15 min
Pickets and Wire Carry – 23 trips carrying objects (5.4 – 15.5kg) at various points along 35m line, following loaded / unloaded pattern. Total of 1.3 km (longest trip of 70.8m).	≤ 17.5 min

As the CMTFE directly reflects the Physical Employment Standard (PES) of the CAF, a separate evaluation, Fitness for Operational Requirements of the CAF Employment (FORCE), was developed to feasibly assess a large sample per annum while balancing adequate fidelity. The FORCE evaluation was designed to ensure that all personnel be capable of performing the essential CMTFE tasks regardless of age, gender, or trade, by capturing relevant movement patterns, energy systems, and muscle group recruitment [4].

> In 2017, the Canadian Army requested a "fitness check" in their Individual Battle Task Standards (IBTS) which are a series of performance goals completed in pre-deployment training with the Canadian Army. FORCE Combat was developed, a modified FORCE test designed to replicate the physiological demands of a dismounted land/urban operation. This evaluation consists of a load bearing 5km march in Battle Order (BO 35 kg) in addition to the FORCE evaluation, in FFO (25 kg), as a continuous event without rest intervals [5]. More information on FORCE Combat can be found at www.forcecombat.ca.

With an established scientific relationship between CMFTE and FORCE, the performance of the six essential tasks can be inferred based on the performance of the four predictive tasks within the FORCE evaluation [4]. This evaluation further assesses health-related fitness, for informational purposes only, through aerobic capacity and waist circumference measurements [4], [6]. CAF personnel who fail the FORCE evaluation two times consecutively are able to register for the CMTFE to fulfil the physical fitness requirement of the universality of service [1], [7]. Table 2-5 highlights the four FORCE tasks, along with respective standards.

Table 2-5: FORCE Evaluation Tasks and Standards [4].

Task	Standard
20 m Rushes – Start prone, 2 shuttle sprints (20 m each way), dropping to prone every 10 m for a total of 80 m.	≤ 51 sec
Sandbag Lift – 30 lifts of a 20 kg sandbag from the floor to 1.0 m mark on wall. Alternates between right and left (1.25 m apart).	≤ 3.5 min
Intermittent Loaded Shuttles – 10 shuttles (20m each way) alternating between load of 20 kg and unloaded, totalling 400m.	≤ 5 min 21 sec
Sandbag Drag – Carry 20 kg sandbag and pull 4 along floor over 20 m.	Pass / fail

2.2.2 Summary of Research to Date to Underpin PES Development

The CMTFE was developed in 2014 through a three phase process, including both objective and subjective methods [8], [9], [10], [11]. Phase 1 focused on the development of an initial list of essential tasks through the review of documents, compilation of literature reviews, post-mission surveys / interviews, and focus group discussions (*Technique for Research of Information by Animation of a Group of Experts – TRIAGE*) [8], [10]. Following a thorough job task analysis process in which 13 essential tasks were identified to be applicable to all environments within CAF, Phase 2 focused on the qualification and quantification of the demands of the tasks. Within individual focus groups consisting of Subject Matter Experts (SMEs) evenly representing all three military environments, Operational Performance Standards (OPS) were reviewed and discussed to generate scenarios in which the tasks are performed [8], [10]. The physical demands of these context-specific scenarios were quantified through biomechanical (e.g., force, load) and physiological (e.g., O_2, consumption, heart rate) measurements. The task list was then reduced with analysis of data from a larger sample guided by a systematic process of "elimination, concentration, and combination" [10]. Tasks with low physical demands were *eliminated;* whereas tasks with elements of low physical demands were *concentrated* to focus on the demanding elements. Tasks were also *combined* when the physical demands (muscle group activation) were analysed to be the same. This process of reduction resulted in a final list of 6 CMTFE tasks [10]. Phase 3 included simulations of the tasks on large, diverse sample sizes to develop a valid and reliable standard test protocol [6], [12]. Special considerations were made for adverse impacts on subgroups, and the impact of technology advancement [10], [12]. Figure 2-2 shows a model of the CAF force evaluation down selection process and their incentivisation tool: the fitness profile.

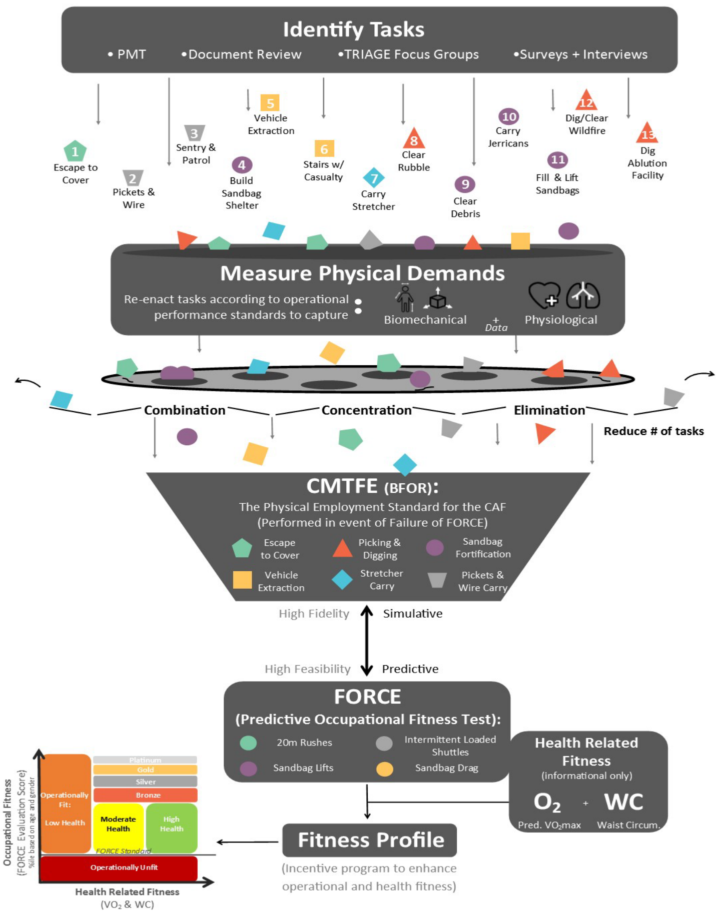


Figure 2-2: A Complete Model of the CAF FORCE Evaluation Down Selection Process and Their Incentivisation Tool: The Fitness Profile.

2.3 UNITED STATES ARMY

Approximate date current PES implemented	January 2017
Approximate duration of the PES	2 h
Test standards the same for gender	Yes
Test standards the same for all ages	Yes
Frequency of PES assessment	Point of selection / When reclassifying to a job with higher physical demands
PES user (e.g., tri-service, single-service or specialist)	Single-service
Is the PES currently being reviewed or redeveloped?	Ongoing evaluation

2.3.1 Summary of Current PES Processes

The Occupational Physical Assessment Test (OPAT) consists of four physical fitness test items: the standing long jump, the seated power throw (medicine ball put), the strength deadlift (hexbar deadlift), and the interval aerobic run (beep test) which can be completed in about 2 hours. The test is taken at the recruiting station and all recruits must pass each test at the required level for their intended job prior to shipping to initial entry training. If a recruit is unable to pass at the level required by their intended job they have the option of retesting, selecting a different job or not enlisting. All jobs are categorized into one of three levels (Heavy, Significant or Moderate). If they are unable to meet the Moderate level, they are advised to train up and retest.

The tests and standards (based on job category) for the OPAT are:

- Standing Long Jump = 120 – 160 cm;
- Seated power throw = 350 – 450 cm;
- Strength Deadlift = 54.5 – 72.7 kg; and
- 20 m Interval Aerobic Run = 36 – 43 Shuttles.

2.3.2 Summary of Research to Date to Underpin PES Development

A job analysis of seven combat arms jobs was conducted to identify and quantify the physiological requirements of the critical tasks of each job [13], [14], [15], [16], [17]. Based on the critical physically demanding tasks, Criterion Measure Task Simulations (CMTSs) were developed, and their reliability assessed [18]. The CMTSs included: casualty evacuation, casualty drag, carry sandbags to build a fighting position, move under direct fire, tactical foot march, loading an Abrams main gun, resupplying an Abrams tank and resupplying a Field Artillery armoured supply vehicle. Lastly, a validation study was conducted that identified four physical fitness tests to screen new recruits into a physically appropriate job [19], [20], [21], [22], [23]. Active duty Soldiers were used to develop these physical fitness tests (named the Occupational Physical Assessment Test, OPAT); however, the intent is to test new recruits. In order to ensure the OPAT would correctly identify new recruits with the potential to perform the physically demanding tasks of their jobs, an additional study was conducted to longitudinally validate the OPAT in new recruits. For that follow-on study, the OPAT was administered at the start of initial entry training and the CMTSs were administered near the end of training. The OPAT correctly identified 76% of soldiers tested into a passing or failing group (passed OPAT and passed CMTSs or failed OPAT and failed CMTSs) [24]. The Soldiers in the validation study are being followed for the first two years of their enlistment to examine attrition, injury, and MOS reclassification. At the point of publication of this technical report, these data are not yet available. The effects of the OPAT on the broader Army population is being assessed quarterly to examine effects on recruiting, retention and injury. (See Section 2.18 for further Background Reading.)

2.4 UNITED STATES MARINE CORPS (USMC)

Approximate date current PES implemented	2015
Approximate duration of the PES	Test-dependent
Test standards the same for gender	Yes
Test standards the same for all ages	Yes
Frequency of PES assessment	Initial Strength Test (IST) – point of selection; Military Occupational Specialty Classification Standards (MCS) – pass out of initial training; Military Occupational Specialty Specific Physical Standards (MSPS) – performed at formal learning center. Sustainment of MSPS occurs at varying intervals.
PES user (e.g., tri-service, single-service or specialist)	Specialist
Is the PES currently being reviewed or redeveloped?	All MSPS standards performance data are routinely reviewed. Biannual assessments occur to ensure quality control of test administration. Periodic feedback is also sought from MOS subject matter experts and operating force personnel to ensure MSPS and associated screening remains relevant and continue to give a reasonable assurance that personnel have the physical capacity to perform in physically demanding occupations.

2.4.1 Summary of Current PES Processes

The USMC refers to its version of Physical Employment Standards (PES) as Military Occupational Specialty Specific Physical Standards (MSPS). There are two types of fitness standards for Marines: general health and physical fitness standards, and MSPS. Physical Fitness Test (PFT) and Combat Fitness Test (CFT) comprise the general health and physical fitness standards. MSPS, implemented on 1 October 2015, are operationally relevant specific physical standards which provide reasonable assurance of success in military occupational specialties. MSPS set the minimum requirements for physically demanding MOSs. MSPS are age- and gender- free and derived from occupationally specific tasks critical for operations (e.g., lift of a 78 lb MK-19 Machine Gun from the deck to overhead, Ground Casualty Evacuation with a 97.3 kg casualty). There are a total of 23 MSPS spread across 29 specified MOSs and both men and women must meet the exact same requirements. They are derived from critical and physically demanding MOS Training and Readiness manual (T&R) tasks and are pass / fail requirements that verify and sustain key physical abilities. MSPS are based on direct individual tasks and surrogate performance for critical crew tasks. They are easily administered within existing Entry Level Training (ELT) Programs Of Instruction (POIs), with remediation opportunities as necessary. MSPS are not the only physical requirements for ELT course completion. Additionally, they are not an assessment of procedural proficiency or fine motor skills, neither are they a pre-requisite for entering an MOS school. Marines who initially do not pass an event are allowed three attempts at passing a MSPS event before being recycled to the next training class. After another three failed attempts, a student Marine is reclassified and assigned another MOS. This allows a total of six opportunities to meet the standard.

Marine recruits and officer candidates take an IST at the point of selection before reporting for ELT which provides an initial physical screening mechanism for those specified MOSs. All Marine recruits and officer candidates are subject to the gender-fair IST standards that are different for males and females. The gender-free IST is required for Marines entering the following Military Occupational Specialties (MOS): Infantry, Artillery, Tanks, Amphibious Vehicles, Combat Engineers, Low Altitude Air Defence, and Ground Ordnance Maintenance. Prior to entering specialty school training, screening is based on performance on specified elements of the PFT and CFT which has been shown in our research to have predictive value in follow-on training.

2.4.2 Summary of Research to Date to Underpin PES Development

MSPS were developed in collaboration with a mix of over 600 officers, Staff Non-Commissioned Officers (SNCO) and Non-Commissioned Officers (NCO) from Training and Education Command (TECOM) including task analysts, schoolhouse instructors, operations officers and chiefs, the Operating Forces (battalion commanding officers, executive officers, and Marines currently serving in the specified MOSs with an average of 6.1 years within their specialty. These individuals comprised the group that participated in job task analyses and Task Criticality Analyses (TCA). The TCA was a critical element of the research which allowed for a rank ordering of tasks from most demanding and most critical to those tasks that were less so. TCA criteria which were weighted based on MOS advocate inputs included task frequency, duration, intensity and importance. MOS advocates were involved in every phase of MSPS development at the Colonel / 0-6 level and higher. Internal Marine Corps organisations involved in MSPS development included Manpower and Reserve Affairs, Counsel for the Commandant (legal review), and the Marine Corps Operational Test and Evaluation Activity. External partners included the Naval Health Research Center, the RAND Corporation, the Government Accounting Office, and the University of Pittsburgh. MSPS standards development began in May 2015 with the testing of over 1,000 Marines from the First Marine Expeditionary Force that served as the basis for MOS specific physical standards recommendations. All personnel tested were current MOS incumbents. The Naval Health Research Center (NHRC) supported study design, data collection, and development of standards methodology. The entirety of MSPS and associated research prior to implementation spanned 2012 to 2015. Follow-on review and evaluation is ongoing.

2.5 UK ROYAL NAVY

Approximate date current PES implemented	Aerobic fitness test (2000), Task simulation test (2013)
Approximate duration of the PES	1 h
Test standards the same for gender	No (aerobic fitness test), Yes (task simulation test)
Test standards the same for all ages	No (aerobic fitness test), Yes (task simulation test)
Frequency of PES assessment	Pre-employment / During Phase 1 training / Annually through career
PES user (e.g., tri-service, single-service or specialist)	Single-service
Is the PES currently being reviewed or redeveloped?	Yes

2.5.1 Summary of Current PES Processes

The in-service Royal Navy Fitness Test (RNFT) consists of two parts: an age and gender-fair field-based aerobic fitness test (Part 1), and an age- and gender-free task simulation test (Part 2). The between-test differences reflect the era of development (*circa* 2000 vs. 2013, respectively). At the pre-employment stage the aerobic fitness test (Part 1) is the only test undertaken and candidates are accepted if they are within 10% of the in-service standard. The in-service standard must be met during Phase 1 training and annually through career. A suite of three test options constitute the aerobic fitness test (Part 1), which includes the, 2.4 km run, 20 m Multi-Stage Fitness Test (ideal for delivery on flight decks or jetties), and the 1.6 km Rockport Walk Test (for those over 40 years of age, or with certain medical exemptions). The tests and standards for the RNFT are:

- Aerobic fitness test (2.4 km run, Multi-Stage Fitness Test, Rockport Walk Test) = 27.6 $ml \cdot kg^{-1} \cdot min^{-1}$ – 46.8 $ml \cdot kg^{-1} \cdot min^{-1}$ (estimated).
- Task Simulation Test = carry × 2, 20 kg powerbags a distance of 60 m in 45 s.

2.5.2 Summary of Research to Date to Underpin PES Development

The origins of both parts of the RNFT derive from the requirements of Royal Navy firefighting tasks. The aerobic fitness standard (Part 1) was developed following an ergonomic [25] and a physical demands analysis of firefighting tasks, which considered both task duration and intensity [26]. The predicted maximal oxygen uptake standard was subsequently adjusted based on age and gender norms [27]. The task simulation test (Part 2) derived from a short-list of tasks that emerged from a subjective job task analysis [28]. Following a job analysis [29] prototype tests were developed and interrogated [30], which resulted in the implementation of an age and gender-free test that simulated the criterion task of firefighting foam drum resupply.

2.6 BRITISH ARMY NON-GROUND CLOSE COMBAT (GCC) ROLES

Approximate date current PES implemented	1998
Approximate duration of the PES	1 hour (pre-employment/selection) and 2 – 3 hours (in-service)
Test standards the same for gender	Yes
Test standards the same for all ages	Yes
Frequency of PES assessment	Point of Selection / Annual
PES user (i.e., generic or specialist	Generic (stratified by job role)
Is the PES currently being reviewed or redeveloped	Yes
If 'Yes', anticipated date when new PES be implemented	2021 for British Army Non-GCC roles

2.6.1 Summary of Current PES Processes

The British Army are currently in the process of a transition for Non-GCC roles from using their existing role-related PES (Physical Selection Standards (Recruits) [PSS(R)] and Annual Fitness Test [AFT]) to a new Non-GCC PES which is due to be completed in 2021. The British Army GCC roles follow the PES process described in the 'UK Armed Forces Ground Close Combat (GCC) Roles' section of the present report.

Applicants to the British Army Non-GCC roles are currently required to perform three gym-based predictor tests; 4 kg Medicine Ball Throw (MBT), Isometric Mid-Thigh Pull (IMTP) and a 2 km run. The pass standards on the predictor tests are related to the physical requirements of the job role to which they are applying to join. The standards originate from the original evidence base for PSS(R), which was established in the 1990s [31], [32], [33], [34], [35] and since implementation in 1998 have been revalidated in 2002 [36] and 2010 [37], [38]. New Standards for Officers [39] and Reservists [40] were latter developed. These entry standards account for changes in physical fitness over the duration of the British Army basic training courses (currently 14 – 48 weeks, depending on role).

The in-service occupational fitness standard for all Non-GCC British Army personnel under the age of 50 years (irrespective of rank) is currently the Annual Fitness Test (AFT). The AFT requires personnel to carry a load of 15, 20 or 25 kg over 8 miles in under 2 hours over varied terrain (a career employment group, gender-free standard).

A series of Operational Fitness Tests (OFTs) which consist of a combination of loaded march tests and mission specific tasks have been developed to assess theatre specific operational fitness for soldiers [41].

2.6.2 Summary of Research to Date to Underpin PES Development

The Job Task Analysis (JTA) which currently underpins PES for the Non-GCC roles in the British Army was conducted in the early 1990s as part of a programme of work to develop PSS(R) [31], [32]. Subjective and objective data gathered in the JTA were to quantify the physical demands of the job tasks performed by personnel. This data was used to inform the development of tests and standards which were implemented in 1998. The PSS(R) tests and standards were then modified during revalidations in 2002 and 2010 as recruiting and training practices developed [36], [37], [38], [39], [40].

The British Army identified that the evidence base underpinning their current PES was becoming outdated and the empirical link between pre-employment, in-service and pre-deployment tests was unclear. Therefore, in July 2015, the Ministry of Defence commissioned research to develop and validate new PES for Ground Close Combat (GCC) roles in the UK Armed Forces. This coincided with a review of whether to lift a ban on women serving in GCC roles (which was lifted in July 2016). The eight GCC roles in the British Army are; Air Assault Infantry (Para), Armoured Regiment, Armoured Cavalry, Armoured Infantry, Mechanised Infantry, Light Cavalry, Light Mechanised Infantry, Light Infantry. The GCC roles in the UK Armed Forces also include; Commando (Royal Marine) and the Royal Air Force (RAF) Regiment. A second research programme was initiated by the British Army in September 2016 to develop new PES for non-GCC roles in the British Army.

The details of the GCC PES research have been reported in the 'UK Armed Forces Ground Close Combat (GCC) Roles' section of the present report. The non-GCC PES research programme is ongoing and is being conducted in four phases in line with the internationally recognised best practice approach for developing PES outlined in the present report, i.e.:

1) Conduct focus groups and surveys to identify the essential physically demanding tasks in each GCC role;
2) Measure the physical requirements and physiological demands of the physically demanding tasks;
3) Develop, and validate, field-based criterion tests based on the physical demands of the physically demanding tasks; and
4) Develop, and validate, pre-employment tests and standards to predict an applicant's likely criterion test results at the end of initial military training.

The research programme is anticipated to deliver pre-employment and in-service PES for the Non-GCC roles by 2021.

2.7 UK ARMED FORCES GROUND CLOSE COMBAT (GCC) ROLES

Approximate date current PES implemented	2019
Approximate duration of the PES	Pre-employment selection: 1 h; In-service; 4 h
Test standards the same for gender	Yes
Test standards the same for all ages	Yes
Frequency of PES assessment	Pre-Employment/Selection and Annually
PES user (i.e., generic or specialist)	Specific for each of the UK Armed Forces GCC roles
Is the PES currently being reviewed or redeveloped	No

2.7.1 Summary of Current PES Processes

The GCC roles in the UK Armed Forces comprise of personnel in the Royal Navy (Royal Marines (RM) Commando (Cdo)), British Army (Air Assault Infantry (Para), Infantry and Royal Armoured Corps) and Royal Air Force (RAF Regiment).

Applicants to the British Army GCC roles and RAF Regiment must pass a pre-employment PES related to their employment role. For the pre-employment PES an applicant's performance is assessed on three predictor tests; 4 kg Medicine Ball Throw (MBT), Isometric Mid-Thigh Pull (IMTP) and a 2 km run. The pass scores on the predictor test are specific to each GCC role and linked to an applicant's probability of passing the In-Service PES for their role at the end of initial training.

After initial training RM, British Army GCC and RAF Regiment personnel are required to annually pass an In-Service PES. The In-Service PES comprises of a series of Representative Military Tasks (RMTs) which simulate the physical demands of the job tasks that personnel are required to perform in their roles. The RMTs are derived from a Job Task Analysis (JTA) conducted with all GCC roles. The RMTs and pass scores vary between the GCC roles:

- RM personnel complete a Rope Climb, Loaded March, Casualty Drag, Water Can Carry and Repeated Lift and Carry;
- Army GCC personnel complete a Loaded March, Fire and Movement, Casualty Drag, Water Can Carry, Vehicle Casualty Extraction and Repeated Lift and Carry; and
- RAF Regiment Personnel complete a Loaded March, Fire and Movement, Casualty Drag, Water Can Carry.

2.7.2 Summary of Research to Date to Underpin PES Development

In July 2015 the Ministry of Defence commissioned research to develop and validate new PES for GCC roles in the UK Armed Forces. Ten GCC roles across the British Armed Forces were examined; Cdo RM, eight GCC roles in the British Army (Air Assault Infantry (Para), Armoured Regiment, Armoured Cavalry, Armoured Infantry, Mechanised Infantry, Light Cavalry, Light Mechanised Infantry, Light Infantry) and RAF Regiment.

The primary data collection was supported by reviews of the national and international civilian and military literature, ongoing stakeholder engagement and collaboration, and Subject Matter Expert (SME) guidance from government organisations, industry professionals and academics. Military Judgement Panels (MJPs) were held during each phase of the research to make evidence-based decisions, led by a single-service senior stakeholder panel of military SMEs.

In Phase 1 of the research programme [42] 10 focus groups attended by a range of ranks (one with each of the 10 GCC roles) and ten rank-stratified sample surveys (covering ≥ 5% of each rank within each of the 10 GCC roles) were undertaken. These were used to identify the most physically demanding tasks performed by personnel in each role. Three single-service MJP1 meetings were conducted to scrutinise the tasks, which resulted in the most critical and physically demanding tasks being down-selected to be observed in Phase 2.

In Phase 2 of the research programme [43] the criterion tasks down-selected by the three-separate single-service MJPs in Phase 1 were observed during 24 training and field exercises conducted by the RMs, British Army and RAF Regt. During these observations, the physical demands of the criterion tasks were documented through notational analysis and measurement of task duration, distances and speeds of movement, mass of equipment, and physical intensity through Heart Rate (HR) and Global Positioning System (GPS) monitoring. The subjective data from the workshops and surveys conducted in Phase 1, and the objective data from Phase 2, were used to complete a Job Task Analysis (JTA) for the 10 GCC roles. The JTA data were used to identify the most physically demanding elements of the criterion tasks and define the task parameters, such as speed of movement, mass of loads and casualties, distances and lifting heights. MJP2 was conducted with each of the RMs, Army and RAF Regt to scrutinise the JTA data and to decide on the most physically demanding elements of the criterion tasks. The MJPs identified the tasks to be measured in Phase 3 to inform the development of Representative Military Tasks (RMTs). The RMTs are single person role-related fitness tests that simulate the physical demands of the job tasks identified in the JTA.

Phase 3 of the research programme was separated into three sub-phases: Phase 3a, 3b [44] and 3c [45]:

1) In Phase 3a, participants from the 10 GCC roles completed Staged Reconstructions (SRs) of the most physically demanding elements of the criterion tasks, which had been agreed at MJP2. The same participants also completed a series of RMTs;
2) The physiological and performance data and the feedback from the participants and military SMEs from Phase 3a were used to refine the RMT protocols to improve the reliability, validity, and ease of future implementation. The revised protocols were endorsed at MJP3 with each of the RMs, Army and RAF Regt to finalise the fixed and performance elements of the RMTs (e.g., distances, speeds, load masses);
3) In Phase 3b, participants performed the revised RMTs (endorsed at MJP3) and a series of predictor tests. The revised RMTs and predictor tasks were performed to an individual best-effort. A subsample of personnel also repeated the RMTs on a second occasion. The participants were from the Cdo (RMs), Army GCC roles, Army non-GCC roles (men and women) and RAF Regt. The Phase 3b data provided information on:
    a) Normative performance,
    b) Test-retest reliability, and
    c) Correlations between RMTs and the predictor tests.
4) A two-stage process was used to agree proposed standards to be evaluated in Phase 3c of the research programme. At Stage 1, a single Tri-Service Standard Setting Workshop (TSSSW) was held with representatives from all three Services (Cdo (RMs), Army, and RAF Regt). At Stage 2, three Single-Service Standard Setting Workshops (SSSSWs) were held; one with each of the services. Additional questions arose from the TSSSW and SSSSW so an Interim Army MJP (MJP3a), and a 16 Air Assault (Air Asslt) Brigade (16X) MJP were also conducted during the process.
5) Phase 3c quantified the pass / fail rate of a representative sample of serving GCC personnel performing the proposed In-Service PES.

At the end of Phase 3 of the research programme three single-service MJP4 meetings confirmed the final combination of RMTs, and associated pass standards to form the In-Service PES for each GCC role [46].

During Phase 4 of the research programme [47] pre-employment PES were developed so that individual's performance on simple gym-based predictor tests could be used to predict the In-Service PES (i.e., RMTs) after accounting for changes in physical performance during training. The participants in Phase 4, consisted of Young Officers (YOs) / Officer Cadets (OCs) and Recruits undertaking training across 14 different training courses; two Royal Marine (RM), eight Army and four RAF Regiment (RAF Regt). Ordinary Least Product (OLP) regression was used to establish the relationships between predictor test and RMT performance to derive an empirical link between the predictor tests and RMTs. These data provided the criterion validity that was considered alongside the content and construct validity to inform the development of the pre-employment PES. The data generated from Phase 4 of the research programme was considered by the RM, Army and RAF Regt at MJP5. Service-specific considerations for the recruitment and training of personnel were also considered at each MJP before agreeing the final pre-employment PES for each GCC role.

2.8 ROYAL AIR FORCE

Approximate date current PES implemented	1994.
Approximate duration of the PES	1 hour
Frequency of PES assessment	At selection, end of Phase 1 training and annually thereafter.
PES user (e.g., tri-service, generic or specialist)	All RAF personnel.
Is the PES currently being reviewed or redeveloped	The RAF Fitness Strategy is under review and PES for roles in addition to RAF Regt will be assessed on an individual Branch and Trade basis.
If 'Yes', anticipated date when new PES be implemented	Ongoing.

2.8.1 Summary of Current PES Processes

All RAF personnel, including the RAF Regt, are required to take the RAF Fitness Test (RAFFT) annually, which measures aerobic capacity and muscular endurance. Aerobic capacity is measured indirectly using the Multi-Stage Fitness Test (MSFT) and muscle endurance is assessed by a 1-min press-up and sit-up test. The RAFFT was introduced to evaluate the effectiveness of the RAF Fitness and Health Strategy, established in 1994 to improve the fitness and health of RAF personnel. Extensive epidemiological research has shown a consistent link between inactivity and illness and premature death [48]. A review of 44 prospective studies has identified a clear dose-response relationship between physical activity and/or aerobic fitness and all-cause mortality [48]. A $\dot{V}O2$ max of 35 $mL.kg^{-1} min^{-1}$ (for males) is the level of aerobic fitness that affords significant protection from disease and mortality [48]. This level was adopted as the minimum aerobic fitness requirement for all male personnel in the RAF, irrespective of age. To ensure that personnel in the RAF who work to age 55 years and beyond can achieve this minimum level of aerobic fitness in later years, aerobic fitness standards were adjusted for younger age groups to take into account the inevitable decline that occurs with ageing [49]. Aerobic standards were set 20% lower for females to take into account the physiological differences between males and females [50]. Test standards for the press-up and sit-up tests were based on population norms standardised for age and gender. A clear link has also been established between relative aerobic fitness and injuries sustained during Phase 1 military training, which justifies different standards for males and females on entry.

2.9 AUSTRALIAN ARMY

Approximate date current PES implemented	2016
Approximate duration of the PES	• All Corps Soldier (baseline) – ~80 min • Combat Arms – ~140 min • Infantry – ~200 min
Frequency of PES assessment	• Basic training – march-out requirement for all recruits • Trade specific training – requirement for some courses, based on employment category (e.g., Combat Arms) and course length • Once every three years in the Combat Brigades
PES user (e.g., tri-service, generic or specialist)	• PES for Australian Army developed and implemented • Three levels; All Corps Soldier (ACS), Combat Arms (CA), Infantry
Is the PES currently being reviewed or redeveloped	The Australian Army PES is neither under review or being redeveloped, however it is under constant surveillance to ensure it achieves organisation intents (e.g., operational capability).
If 'Yes', anticipated date when new PES be implemented	N/A

2.9.1 Summary of Current PES Processes

The development of the Australian Army PES was based on the 'soldier first' concept, i.e., irrespective of occupational specialty; all personnel are expected to be able to perform generic tasks related to military service (Table 2-6). These generic military tasks provided a reference point for all employment category-specific tasks and demands, and the subsequent development of PES (Table 2-7 and Table 2-8).

All Australian Army personnel are required to meet PES at various junctures in their career. Both recruits and officer cadets are required to meet PES standards prior to graduating to the incumbent workforce. The incumbent workforce within the Combat Brigades is required to achieve PES every three years based on the force generation cycle. The force generation cycle involves annual phases; readying, ready, reset. At the end of readying phase, and prior to being 'ready' (i.e., ready for deployment), personnel within the respective combat brigade are required to achieve their employment category-specific PES.

Table 2-6: Australian Army PES Components.

Forced March	The Forced March tests the ability to undertake load carriage (e.g., patrolling, administrative movements). The aerobic demands associated with load carriage also have relevance to tasks such as digging, sand-bagging and other repetitive manual handling tasks.
Fire and Movement	The Fire and Movement assessment tests the ability of personnel to perform a loaded, high intensity, short duration task that may be expected when operating in either an offensive (e.g., fire and movement) or defensive (break contact) environment.
Lift and Carry	The Lift and Carry tests local muscular endurance. It is based on the requirements of performing a stretcher carry, but has relevance to other muscular endurance tasks such as admin resupply and carrying defence stores.

Box Lift	The Box Lift (BL) is a task simulation linked to the requirement to lift a field pack into the back of a common military vehicle. The assessment is designed to evaluate a soldier's ability to perform manual material handling whilst maintaining correct lifting technique.

Some employment categories in Army have higher physical demands than others and as such require more demanding assessments. The routine tasks of each employment category are aligned to one of four PES Levels (see Table 2-7).

Table 2-7: Four PES Levels for Each Employment Category.

Level 1	All Corps Standard (ACS)	The ACS PES is based on the performance of a range of basic military tasks including preparing defensive positions, local patrolling, fire and movement and conducting casualty evacuation.
Level 2	Combat Arms (CA)	The CA PES is based on the higher physical demands of conducting combat operations, or operating in a high threat environment. This baseline is applied to most combat arms employment categories (Artillery, Armoured, Combat Engineers), but may also be applied for specific ACS roles in support of combat operations.
Level 3	Higher Demands (HD)	Level 3 is based on the higher physical demands across all components of PES (Infantry).
Level 4	Increased BL	Level 4 is based on increased physical demands for muscular strength (Artillery).

Table 2-8: Specific PES Requirements.

	Forced March	**Fire and Movement**	**Lift and Carry**	**Box Lift**
Overview	March at a rate of 5.5 km/h	6 m bounds (every 20 sec) to a cadence	Carry two 22 kg jerry cans or 22 kg kettle bells in 25 m legs to a cadence	Lift a weighted box from the ground to shoulder height using a prescribed lifting technique
Level 1	5 km with a 20 – 23 kg load within a 50 – 55 min	12 x 6 m bounds	5 x 25m legs	25 kg
Level 2	10 km with a 35 – 40 kg load within 100 – 110 min	16 x 6 m bounds and 18 m leopard crawl	11 x 25 m legs	30 kg
Level 3	15 km with a 4 0 – 45 kg load within 150 – 165 min	1 km move in 8 mins, 16 x 6 m bounds and 18 m leopard crawl	11 x 25 m legs	35 kg
Level 4	n/a	n/a	n/a	40 kg
Refs	[51], [52]	[53], [54], [55]	[56], [57], [58], [59]	[60], [61], [62], [63], [64], [65]

2.9.2 Summary of Research to Date to Underpin PES Development

The development of PES assessments and standards involved extensive field observations and workshops with relevant Army personnel across all employment categories. All identified physically demanding tasks were grouped by researchers based on the physical capacity (muscular strength, muscular endurance, aerobic capacity, and anaerobic capacity) identified most likely to limit task performance. Beyond 'trade' specific tasks, it was determined that all Army personnel were required to perform generic 'soldiering' tasks. These tasks included load carriage, fire and movement, casualty evacuation and manual handling, and subsequently formed the basis for Army PES. (See Section 2.18 for further Background Reading.)

2.10 NEW ZEALAND ARMY

Approximate date current PES implemented	2015
Approximate duration of the PES	12 persons in a 1-hour period (36 in 90 min)
Test standards the same for gender	Yes
Test standards the same for all ages	Yes
Frequency of PES assessment	Annual
PES user (e.g., tri-service, generic or specialist)	All NZ Army Regular Force soldiers
Is the PES currently being reviewed or redeveloped	No
If 'Yes', anticipated date when new PES be implemented	NA

2.10.1 Summary of Current PES Processes

The New Zealand Army Land Combat Fitness Test (LCFT) was designed to assess an individual's physical readiness or otherwise to perform the required tasks at D-LOC (Directed Level of Operational Capability). The test is performed annually and typically takes one hour to complete. The tasks being simulated are a tactical movement, a fire and maneuver, a lift and place and a lift and carry. If an individual fails any component of the test, they fail the assessment and are required to be retested within four weeks. If a second failure occurs, specific remedial training is prescribed by NZ Army PTIs in direct consultation with soldier and command. Subsequent LCFT testing is on SME advice from PTIs and at the discretion/direction of command.

2.10.2 Summary of Research to Date to Underpin PES Development

All physical tasks performed by the All Arms Soldier, regardless of trade, gender, age or training, were identified via extensive focus group discussions, SME input and review of doctrinal evidence [66]. Identified tasks were then classified and ranked to determine the most physically demanding within each of four physical component categories – aerobic power, anaerobic power, muscular strength and muscular endurance. On completion of this process, the following five physical tasks / duties were confirmed as being both the most physically demanding and common to every All Arms Soldier.

1) Extended load carriage (tactical movement) – a 4 km tactical movement was considered a representative military task. Rate of movement would be faster than a fast forced march speed, but slower than slow running pace. It was considered acceptable that soldiers perform this activity in a 50/50 walk/run fashion and as such, an average minimum movement rate of the mid-point of these two speeds (7.5 $km.h^{-1}$), was considered appropriate.

2) Fire and manoeuvre (F&M) over a distance of 150 metres in high-density bush/obstructions, with an exposure time of no more than 5 s per bound. This was found to represent the most physically demanding but realistic example of F&M [67].
3) CASEVAC extraction / body drag in a team of two, pulling the casualty backwards by his/her webbing over a distance of 15 metres to an initial point of safety. A body drag was removed from the final test battery as early analysis revealed that soldiers who could perform Lift and Place and Lift and Carry activities could also perform the body drag, but not vice versa.
4) Lifting ten filled jerry cans (20 kg) on to and off the deck of a Unimog vehicle, in a continuous fashion.
5) Carrying a stretcher, in a team of 2, over a distance of 200 metres, at a rate of 1.25 $m.s^{-1}$ (4.5 $km.h^{-1}$).

Based on the identified criterion tasks, a pilot test battery of simulation tasks was developed. The tasks were renamed and reordered to reflect the physical demand and common terminology:

1) **Lift and Place (L&C).** Lift and place a 20 kg jerry can on to the deck of a vehicle ~1.4 m high. Lower to ground and repeat 10 times in 90 seconds.
2) **Battlefield Manoeuvre (BM).** Advance 150 metres in 15 x 10 metre bounds, each completed within 5 s with a 5 s rest between.
3) **Lift and Carry (L&C).** Lift and carry 2 x 20 kg jerry cans over a distance of 200 metres, in 8 x 25 metre stages. Each 25 metre stage to be completed within 20 s, with a 5 s rest between.
4) **Battlefield Endurance (BE).** Complete a 4 km movement in less than 32 minutes but no less than 30 minutes in order that subsequent tasks could be performed if required.

Test criteria and standards represent the minimum acceptable level for an All Arms Soldier and are pass or fail. The order of tests ensures that performance on one test component should not compromise performance on the next. Body armour, webbing and weapon with a combined weight of 20 kg are required to be worn/carried throughout the test to reflect a minimum, mission appropriate, fighting order load.

To assess test-retest reliability, 24 soldiers performed the LCFT on two occasions, 72 hours apart. The Coefficient of Variation (CV) was 5.5% which compares favourably to that seen for other physical performance tests and indicates acceptable reliability [68].

The test battery was piloted with 1090 soldiers who performed 1159 LCFTs during a 10 month period. In total, there were 1034 passes and 125 failures (94 individuals). Failure rate was highest for the Battlefield Endurance component of the test and lowest for the Lift and Place, regardless of gender. A higher proportion of females than males failed the test. Of the 94 individuals who failed the initial test, 43 attempted the test again and over half were successful at the second attempt. The vast majority of soldiers considered that all components of the test were relevant to the All Arms Soldier, but at the time, fewer considered scheduled PT at the time to reflect the physical demand of the tasks.

2.11 ROYAL NEW ZEALAND AIR FORCE

Approximate date current PES implemented	2006
Approximate duration of the PES	30 persons in 90 minutes
Test standards the same for gender	No
Test standards the same for all ages	No

Frequency of PES assessment	6 monthly for a standard pass; annually for an excellent pass
PES user (e.g., tri-service, generic or specialist)	All RNZAF
Is the PES currently being reviewed or redeveloped	No but this has been recommended
If 'Yes', anticipated date when new PES be implemented	NA

2.11.1 Summary of Current PES Processes

The aim of the RNZAF Operational Fitness Test (OFT) is to confirm individual physical fitness for operational deployment and is also used to determine eligibility for promotion, some formal training, and participation in some sporting activities.

Part 1 of the test is a push-up test, with different standards based on age and gender.

Part 2 of the test is a timed 5 km march carrying an evenly distributed weight of 20 kg, which is representative of a helmet, flak jacket, Steyr, 120 rounds of ammunition, webbing, 1.5L water, and a first aid kit. Time allowed to complete this task is different for males and females of different ages. If an individual fails either part of the test, they fail the assessment. They are then given 28 days to pass and an initial warning. If the individual does not pass within 28 days they are given a formal warning, placed on remedial training and must pass within the next 90 days. After 90 days the individual's Commanding Officer must raise a recommendation concerning the retention of the individual in the RNZAF.

2.11.2 Summary of Research to Date to Underpin PES Development

In 2002, a project team was established to look at a new Operational Fitness Test for the RNZAF. The remit for the team was to ensure that the resulting test:

1) Reflects the physical requirements that could be expected under deployed conditions and is therefore viewed by all personnel as an acceptable pre-requisite level of personnel fitness for deployment;
2) Is challenging but not impossible;
3) Is pass/fail;
4) Is not gender- or age-biased;
5) Will detect physical fitness deficiencies in unfit personnel;
6) Is easily administered utilising the minimum PTI personnel;
7) Is standardised across the bases; and
8) Is safe.

In developing the test, the project team considered roles undertaken in deployments such as East Timor and those identified by the Royal Australian Air Force. McAra [69] highlighted that no adjustments should be made for age and gender as everybody is expected, in addition to their primary duties, to have the capacity to complete the following tasks:

- Erecting large tents;
- Filling sandbags;

- Building defensive barricades;
- Digging foxholes;
- Constructing outdoor latrines and showers;
- Loading and unloading supplies from vehicles and aircraft;
- Erecting radio antennae, power poles, camouflage nets;
- Stretcher bearing;
- Patrolling;
- Working long hours on flight line repairing aircraft; and
- Erecting barbed wire.

The project team concluded that Strength and Endurance were "Essential" attributes, whilst Agility, Speed and Coordination were "Desirable" attributes. When reviewing fitness testing options, the team looked to the RAAF and firefighters tests which included a 5 km march to assess endurance. A completion time of 46:00 minutes was selected based on the tests of other militaries and emergency services and a trial of 200 randomly selected serving personnel. The 20 kg weight to be carried was equivalent to the patrol weight in a high threat environment such as East Timor. A push-up test was selected as an easy to administer indicator of upper body strength. A score of 12 was selected as the standard, based on that used by other militaries and a trial of serving personnel. Completion of the 5 km endurance test in a time of 43 minutes or less, along with a push-up score of 25 or more was considered an Excellent pass. The McAra [69] report recommended that this new test, the RNZAF Operational Fitness Test (OFT), be trialled for a 12 month period to confirm its suitability, but a review of this trial was not completed. In 2006, a further review of Physical Fitness in the RNZAF was conducted [70] 2006) which emphasised that a high standard of physical fitness was required to i) ensure that personnel can rapidly acclimatize to different environments and achieve the physical requirements that can reasonably be expected under a range of deployed conditions; ii) maintain the general health and wellbeing of personnel to both improve performance in the workplace and reduce the health care costs incurred by the RNZAF; and iii) portray a positive public image of the RNZAF. As a result of this review, a new RNZAF Fitness Policy for Regular Force (RF) personnel was approved and the Operational Fitness Test (OFT) formally introduced. A Restricted Fitness Test was also created, for personnel with long-term injury or illness. Unlike the originally proposed OFT, the implemented OFT had different standards based on age and gender and there is no record of the decision to enforce this.

In 2017, a review of fitness testing across the New Zealand Defence Force [71] highlighted that push-up score has not been proven to relate to job performance and that different standards for the 5 km loaded march are not justified. A full review of the RNZAF OFT and standards has been recommended.

2.12 ROYAL NEW ZEALAND NAVY

Armed Forces Branch	Royal New Zealand Navy
Approximate date current PES implemented	2015
Approximate duration of the PES	12 persons in a 30 minute period
Test standards the same for gender	Yes, for job-related component
Test standards the same for all ages	Yes, for job-related component
Frequency of PES assessment	6 monthly
PES user (e.g., tri-service, generic or specialist)	All sailors

Is the PES currently being reviewed or redeveloped	No
If 'Yes', anticipated date when new PES be implemented	NA

2.12.1 Summary of Current PES Processes

The Royal New Zealand Navy fitness test is performed annually and typically takes 30 minutes to complete. The test has two components – a health-related fitness component (the Multi-Stage Fitness Test) and a job-related fitness component (a lift and carry and a dummy drag). If an individual fails any aspect of the test, they fail the assessment. They are then given 4 weeks to pass.

2.12.2 Summary of Research to Date to Underpin PES Development

Focus group discussion and one-on-one interviews with sailors, Command, training instructors and the Maritime Operational Evaluation Team (MOET) were initiated to identify the most physically demanding tasks expected of all seagoing personnel [72]. The following tasks were identified and subsequently observed and measured:

1) Lift and carry of an AFFF container – weight carried was 20 kg over 20 metres.
2) Individual body drag out of a compartment – mean mass of an RNZ sailor is 86 kg and maximum distance from a bunk to the outside of a compartment is 15 metres.

Damage control (DC) was discussed at length but not taken forward for subsequent analysis as:

a) Subject Matter Experts (SMEs) determined that the task is not performed by all personnel;
b) Muscular strength and endurance requirements are no greater than those seen for the other identified tasks; and
c) Aerobic demand has been previously found to be low [73].

However, personnel are often required to work extended hours, in challenging environments, where a low level of aerobic fitness may impact performance, resilience, injury risk and long-term health. A requirement for all personnel to have at least an 'average' level of aerobic fitness for their age and gender is therefore justified.

Based on the above, the following tests were selected and implemented to assess the identified physical competence and health-related fitness requirements:

a) Lift and carry of a 20 kg kettle bell (4 x 15 metres) in 45 seconds.
b) Weighted mannequin drag (86 kg) over 15 metres in 30 seconds.
c) Multi-Stage Fitness Test (MSFT).

To reflect the benefits of all seagoing personnel being able to perform DC team activities, the research team recommended that formal assessment of this capability be introduced during pre-deployment refresher sea survival training.

Based on the known relationship between aerobic fitness and all-cause mortality, a requirement for all personnel to have at least an 'average' level of aerobic fitness for their age and gender was justified, although strategies to encourage exceeding this level are being explored to reflect the benefits of improved health and stamina for operational effectiveness and long-term quality of life. 'Average' aerobic fitness was defined based on the work of Kodama et al. [74], and the Canadian Defence Force [75]. Based on this evidence, the age and gender-fair recommended minimum MSFT scores for RNZN personnel are presented in Table 2-9 as Level and Shuttle, with the associated VO_{2max} in $ml.kg^{-1} \cdot min^{-1}$ in brackets.

Table 2-9: Age- and Gender-Fair Recommended Minimum MSFT Scores for RNZN Personnel Presented as Level and Shuttle, (VO_{2max} in $ml.kg^{-1}.min^{-1}$).

	< 30 years	30 – 39 years	40+ years
Males	7.10 (39.9)	6.10 (36.4)	5.09 (32.9)
Females	5.09 (32.9)	4.09 (29.5)	4.01 (25.9)

2.13 BUNDESWEHR

Approximate date current PES implemented	2010
Approximate duration of the PES	90 minutes (Basic PES)
Test standards the same for gender	Yes
Test standards the same for all ages	Yes
Frequency of PES assessment	Point of selection / Annual
PES user (e.g., tri-service, single-service or specialist)	Generic
Is the PES currently being reviewed or redeveloped	No
If 'Yes', anticipated date when new PES be implemented	N/A

2.13.1 Summary of Current PES Processes

The Bundeswehr defines military fitness as a three-level construct:

1) "Fundamental/Baseline-Fitness",
2) "Basic-Military-Fitness", and
3) "Task Fitness" in the "Directive: Sport and Physical Performance".

The directive establishes standard guidelines for regular physical performance assessments of the first two levels (Fundamental/Baseline-Fitness and Basic-Military-Fitness). These two levels concern all active soldiers of all services (Joint Forces Level) regardless of age, gender, rank, and occupation.

1) **"Fundamental/Baseline-Fitness".** Since 2010 "Fundamental/Baseline-Fitness" is assessed once a year with the Basic Fitness Test (BFT) [76], [77], [78]. It consists of three tasks in a track suit that have to be completed within 90 minutes in the following order:
   a) 11 x 10 m shuttle run (≤ 60 sec);
   b) Flexed arm hang in the chin up position (≥ 5 sec); and
   c) 1000 m run (≤ 6:30 min).

The BFT is also part of the recruiting process at the assessment centers of the Bundeswehr. Every applicant (temporary-career volunteer) has to pass the tests as specified above, the only difference is that the 1000 m run is replaced by a bicycle ergometer test where the minimum standard is 06:30 min at 130 W and 80 rpm. Voluntary service conscripts (≤ 23 months of voluntary service) do not carry out the BFT during the recruiting process.

2) **"Basic-Military-Fitness".** Every soldier has to pass the following two tests every year (times, distances, weights, and minimum criteria are identical for all genders and ages):

   a) Ruck-march in field uniform (no helmet, no ballistic body armour) with a 15 kg backpack. The minimum criteria for passing for all soldiers is to cover 6 km in 60 min. In order to qualify for the "German Armed Forces Badge for Military Proficiency" criteria are 9 km in 90 min (Silver) or 12 km in 120 min (Gold).

   b) 100 m swimming fully dressed (swimming suit + field uniform pants + field uniform jacket) in 4 minutes with subsequent undressing of the field uniform in the water or 200m swimming (including a jumping start) in 7 minutes.

For a more comprehensive assessment these tests are complemented with the "Basic-Military-Fitness-Tool" (BMFT) [79], [80], [81], [82], [83], [84]. It was developed to be the force-wide monitoring tool for both pre-deployment training and military fitness training in general. The BMFT consists of the following tasks in field uniform with helmet and ballistic vest:

- Maneuver under fire;
- Casualty rescue;
- Load carrying; and
- Load lifting.

3) **"Task Fitness".** Currently no tests for the assessment of task fitness have been implemented. In the future, occupation-specific tests, for example for Infantry or Engineer troops, will be developed to assess job-related proficiency levels. The responsibility for the definition of tests and corresponding fitness levels lies with the different services and branches.

2.13.1.1 German Sports Badge of the German Olympic Sports Confederation

Soldiers also have to pass the standards of the German Sports Badge of the German Olympic Sports Confederation annually.

2.13.1.2 Military Medical Assessment

During his or her active service period every soldier of the Bundeswehr has to undergo different forms of medical examinations directly tied to specific occasions (e.g., deployment, change of career or status) or occupational specialties (e.g., fitness for military flying duties or shipboard duty). As of January 2019 the Bundeswehr has implemented a new system for a regular medical assessment (General Application Examination – Individual Basic Skills = "AVU-IGF") in a three year interval. It is compulsory for every soldier and intended to ensure medical readiness for the yearly testing of the Individual Basic Skills (self and buddy aid, elementary NBC protective measures, and marksmanship) and the specified sport and fitness tests. The results are also intended provide information for individual preventive medical counselling. Prospectively, the aggregated data of the AVU-IGF results will be used to establish a health-oriented operational situation report for the Bundeswehr command and serve as data base for health promotion and prevention. Assessments of medical aptitude will continue for specific occasions (career change, occupational specialties). However, neither of the exams within the framework of military medical assessment in the Bundeswehr take physical performance into account. To close this gap, a research program for the development of a reliable instrument for categorization of physical performance was initiated, in 2017.

2.13.2 Summary of Research to Date to Underpin PES Development

2.13.2.1 Fundamental/Baseline-Fitness

The Basic Fitness Test (BFT) was introduced in 2010. It monitors fitness components relevant for military duty (strength, endurance, speed/agility) derived from on-site task analyses. The BFT has to be completed once a year by every soldier. Results are adjusted based on age and gender. The BFT rating system was reevaluated and readjusted in 2014 [76], [77], [78], [85]. Currently, BFT results are only used as criteria during the recruitment process. BFT performance does not play a role in personnel selection or assignment within the active duty members of the Bundeswehr.

2.13.2.2 Basic Military-Fitness

The Bundeswehr currently employs the following three means to assess and monitor Basic-Military-Fitness:

- 6 km loaded march (15kg);
- 100 m swimming fully dressed in 4 minutes or 200m swimming (including a jumping start) in 7 minutes; and
- The BMFT (Basic-Military-Fitness-Tool) for assessing "Basic-Military-Fitness" during pre-deployment training and the basic training of the German Army. The BMFT was developed and evaluated in two separate studies [76], [80], [81], [83], [84]. The fundamental pre-requisite was to compile a physical requirement profile reflecting physical stress components experienced in deployments at a joint forces level. Mission-typical tasks were identified in targeted on-site analyses of different branches during pre-deployment training at the Army Combat Maneuver Training Center. After conducting further analyses and incorporating lessons learned from operations, the operational requirements as regards exercise physiology with accentuation of the aspects movement and load were operationalized in a level-appropriate BMFT. The field uniform test "Basic-Military-Fitness-Tool" contains four tasks:

  a) Maneuver under fire;

  b) Casualty rescue;

  c) Load carrying; and

  d) Load lifting [76], [80], [81], [83], [84].

  Like the BFT, the BMFT is currently not used for assignment of military personnel.

2.13.2.3 Military Medical Assessment

There is no comprehensive instrument for the assessment of physical performance within the framework of military medical assessment the Bundeswehr. Therefore, a research program was initiated in 2017, in which a reliable instrument for the categorization of physical performance will be developed. This instrument is intended to be the standardised assessment method during military medical examinations.

2.14 FRENCH ARMED FORCES

Approximate date current PES implemented	Not yet implemented
Approximate duration of the PES –	General physical fitness (CCPG) 4 h and Specific physical fitness test (CCPS) – 4 h

Test standards the same for gender	Same test but with different standards for men and women
Test standards the same for all ages	Same test but with different standards for older and younger
Frequency of PES assessment	Annual
PES user (e.g., tri-service, single-service or specialist)	Generic
Is the PES currently being reviewed or redeveloped	The French Army is currently considering a new PES to assess the capability to fight with a load.
If 'Yes', anticipated date when new PES be implemented	2020

2.14.1 Summary of Current PES Processes

In France, assessment of the military combat fitness (CCPM) is divided in two parts. One concerns general physical fitness (CCPG), compulsory and common to all Armed Forces and the other consists of specific test (CCPS), optional and decided by each service but which does not evaluate the ability to perform the mission or the position.

The CCPG is proposed by the Chief of the Defence Staff and practical application are defined by each Army chief. This evaluation is annual and the scoring scale depends on age and sex. CCPG counts for 60 points and takes half a day. The aim is to assess a fitness level but in fact only used to mark military personnel and not really to better guide the role and employment of personnel. The current tests are thought to simulate or predict a task. Currently, no specific standard is required before deployment. The CCPG assesses three physical fitness elements:

1) Cardiorespiratory Endurance (ECR);
2) Aquatic Ease (AA); and
3) General Muscle Fitness (CMG).

Each group is graded with 20 points, thus a score of 60 points can be scored for all three.

Three modalities are available to assess the Cardiorespiratory Endurance (ECR). (All performed wearing running shoes):

1) **ECR 1: Running of 12 minutes named Cooper test.** Consists of 12 minutes running as fast as possible, on track and field.
2) **ECR 2: VAMEVAL.** An evaluation of maximal aerobic speed on a track and field marked each 20 metres. First pace of running is 8.5 km/h and pace increases regularly by 0.5km/h each minute until subject is unable to follow the pace.
3) **ECR 3: Shuttle test of 20 m (Luc Léger).** A 20 m shuttle run with increment of 0.5 km/h each minute beginning at 9 km/h.

The Aquatic Ease (AA) consists of a jump in water, a 100 m swim whatever the style immediately followed by a 10 m breath hold.

Three modalities are available to assess the General Muscle Fitness (CMG). The choice of one or another is free and determined by each Army. In each combination, each exercise is awarded 10 points.

1) **CMG 1**

   a) **Rope climbing (*corde* in French):** Climb a 5 m smooth rope whatever the style, twice successively which corresponds to an actual 7 m climb.

   b) **Sit-up (*abdominaux* in French)** Consists of performing as many sit-ups as possible in 2 minutes.

2) **CMG 2**

   a) **Push-up (*pompes* in French).**

   For men – Bending the upper limb up to 20 cm from the ground.

   For women – Bending the upper limb up to 20 cm from the ground.

   b) And **Sit-up (**as above).

3) **CMG 3**

   a) **Pull-up (*tractions* in French).** Upper limb in pronation, at a distant of shoulders width. Performing as many repetitions as possible without limit of time.

   b) **And Sit-up** (as above).

From this general physical fitness physical test, the French Army has chosen among the different modalities mentioned above:

- For ECR, the first modality that is the Cooper test.
- For CMG, the first combination that is rope climbing and sit-up described above.

The French Army have also added some specific tests (CCPS) counting for 40 points and takes a half day. They have been selected to assess its specific fitness needs; a long run in army boots and a shoot.

Walk and Run (20 points) consists of covering 6 to 8 km depending on age, on flat field in army boots without external load, nor weapon in short a time as possible.

Shooting (20 points) consists of a shoot of 10 ammunitions FAMAS rifle. Five targets can be touched with 2 cartridges per position. The 2 remaining cartridge can be used in case of failure at a position and without penalty. The four positions are mentioned below:

- Prone at 100 m from the target, in less than 6 seconds;
- Kneeling down or squatting , at 75 m form the target, in less than 5 seconds;
- Standing up at 50 m in less than 5 seconds; and
- Standing up at 25 m, in less than 5 seconds, twice.

Four points are awarded per target reached. If the shooter runs out of time, the shoot to this position is awarded a score of 0.

Depending results from CCPG, five levels of physical fitness have been defined as shown in Figure 2-3.

CLASSEMENT CCPG				
5	4	3	2	1
60 51	50........ 41	40........ 31	30 21	20...... 0

Figure 2-3: The Levels of Fitness as Defined by the CCPG (i.e., General Physical Fitness) Classifications.

According to physical fitness level, three groups of demanding tasks have been identified requiring low, medium or high physical fitness (see Figure 2-4). But each position is not described and the different jobs not classified in the three groups. Thus position is decided by command and not according to PES or physical fitness level.

If a person has less than 26 points, a personal aim of improvement must be given to achieve the level required in three or six months.

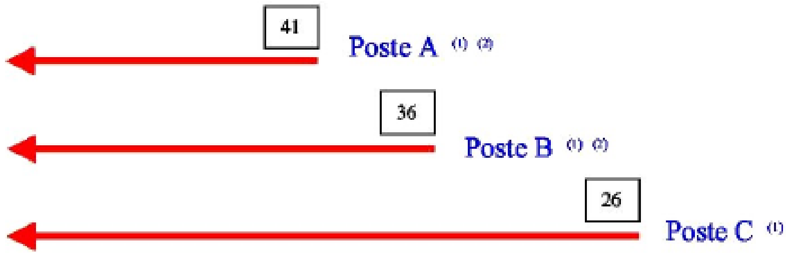


Figure 2-4: The Three Levels of Physical Fitness Decided by Command that Personnel are Required to Achieve.

2.14.1.1 Parachute Regiment

A 1500 m walk and run is performed with a back pack weighting 11 kg for men and 5 kg for women in a time under 9 minutes. In addition, men must perform at least 4 pulls-up and women must maintain a hang for 30 s.

Moreover, it certain anthropometrics standards exist related to the type of parachute used:

- Height < 195 cm AND 55 kg < body weight < 90 kg.
- Upper 90 kg, body weight must be > 110 kg AND height < 2.05 m AND BMI <26.5.
- Upper 195 cm, 20<BMI <26.5 AND height <2.05 m.

BMI should not be upper 25 kg/m^2. For very strong subjects, a BMI > 27 is accepted and abdominal perimeter must be under 94 cm for men and 80 cm for women.

2.14.1.2 Mountain Troop

Instead of the 8 km walk and run, the Mountain Troop must also complete the Military skier's certificate (BSM) and Military mountaineering certificate or (BAM). This consists in a 1000 – 1500 m ascension of positive slope (around 10 km length) with a back pack of 14 kg for men and 8 kg for women with riffle to perform in a minimal time depending on the course (3H30 to 5H00).

2.14.2 Summary of Research to Date to Underpin PES Development:

The Infantry school is going to implement another specific physical test that assesses readiness to fight with a load. This is the first test generated from basic infantry tasks and job of an infantryman. This is a Load Effect Assessment Program, called Test de Puissance du Combatant (TPC), is a loaded obstacle course. It consists of wearing the equipment of the dismounted soldier (loaded 37 kg with rifle, helmet, backpack) to perform a 500 m circuit:

- With obstacles: cross low or mid-high but never high obstacles;
- Repeated shooting positions;
- Drag a wounded soldier of 100 kg; and
- Translate a load of 20 kg on 20 m.

To succeed in the test, the individual should not fail their shoot and perform the course in a minimal time. The threshold would be gender- and age-free. This proposal from infantry has not yet validated by the French Army.

Currently, each Army branch are considering new specific tests more in accordance with the job tasks which would be gender-free and with a minimal level to pass.

Concomitantly, the CCPG test would be simplified to only assess physical fitness with a running test, push-up and aquatic ease. Include each year in the marking of military people, CCPG would stay with different standards for men and women. (See Section 2.18 for further Background Reading.)

2.15 DANISH ARMED FORCES

Approximate date current PES implemented	2017
Approximate duration of the PES	45 – 60 min
Test standards the same for gender	Yes
Test standards the same for all ages	No – there are three age categories with lower pass standards for older age
Frequency of PES assessment	Annual
PES user (e.g., tri-service, single-service or specialist)	Tri-service
Is the PES currently being reviewed or redeveloped	No

2.15.1 Summary of Current PES Processes

All military personnel in Denmark have had to pass a PES called the Basic Physical Fitness Test (BPFT) every year to fulfil the demands of employment within the Danish Defence Force. It is the same test whether they are in the Army, Navy or Air force. The purpose of the BPFT is to measure a physical minimum requirement (termed CATEGORY 1) of the Operational Physical Requirements (OPR). The OPR describe the physical recommendations for any given task/job in the Danish Armed Forces and reflect the ability to handle intense workloads when deployed nationally or internationally. The requirements of the BPFT and the OPR are illustrated in Table 2-10. The basic fitness test is differentiated on age but not gender, the OPR are not differentiated at all. Both the BPFT and OPR are divided in to two elements; Test A and Test B: Test A measures endurance to measure the ability to recover from strenuous physical exercise/work and a measure of the ability to acclimatize. Depending on the practical feasibility either a 12 min run or shuttle run tests can be used. Test B measures aerobic and anaerobic intermittent work to measure the ability to perform repeated high intensity work. The movement pattern reflects the peak performance work demands for soldiers e.g., performing bounds when in contact with the enemy or conducting firefighting or damage control on a ship. The Danish Military Speed Test (DMST), requires the soldiers to run back and forth between two lines, 20m apart, for as many rounds as possible in 30 seconds, followed by 30 seconds of rest. The cycle is repeated ten times. The completed number of rounds is summed up and counts as the score in the test. The BPFT is the measurement of CAT 1 and the OPR covers four further categories:

- CAT 2: Functions with low physical demands e.g., Administrative work.
- CAT 3: Functions with moderate physical demands e.g., Guard duty or mechanics.
- CAT 4: Functions with predominantly moderate physical demands e.g., “Gunner” or truck drivers.
- CAT 5: Functions with predominantly high physical demands e.g., Infantry or security forces.

The Basic Physical Requirement (CAT 1). If the test is failed the commander prepares, together with the military physical trainer / advisor an 8 – 12 week training program after which the test is carried out again. If the test still isn't passed, actions will be taken in order to be able to decide about continued employment or dismissal from the Danish Armed Forces.

Operative Physical Requirements (CAT 2-5). All operative functions are placed in category 2-5 according to the physical requirements of the function. Cat 5 includes operative functions with the highest physical requirements. Hence there is an increase in the physical requirements from CAT 2 to CAT 5. All military personnel have to pass OPR, according to their operative function (CAT 2-5), when deployed nationally or internationally. If not, the commander prepares, together with the military physical trainer / advisor an 8 – 12 week training program after which the OPR is carried out again. If the test still isn't passed, the commander has the final decision concerning participation in the current deployment.

Table 2-10: The Requirements of the Basic Physical Fitness Test and Operational Physical Requirements (OPR).

DANISH DEFENCE PHYSICAL REQUIREMENTS

BASIC TEST		ENDURANCE*			STRENGTH				
CATEGORY / AGE		12 MIN	20 m Shuttle Run		SPLIT SQUAT	DIPS ON A BENCH	HORIZONTAL PULL-UP		BURPEES
		Metre	Level		Set x Reps Per Leg	Set x Reps	Set x Reps		Reps
CAT 1	60 –	1800	5.9		3 x 10	3 x 10	3 x 2		5
	50 – 59	2000	6.9		3 x 12	3 x 12	3 x 3		10
	– 49	2200	8.3		3 x 15	3 x 15	3 x 5		15

OPERATIVE TEST	TEST A*		TEST B	STRENGTH				
CATEGORY	12 MIN	BIP TEST	DMST	LUNGES	DIPS	PULL-UP	DEADLIFT	BURPEES
	Meter	Level	Laps	10 Reps Per Leg	Reps + kg	Reps + kg	5 Reps	Reps in 5 min.
CAT 2	2400	9.6	54	20 kg	1 + 0 kg	1 + 0 kg	40 kg	35
CAT 3	2500	10.2	56	35 kg	3 + 0 kg	3 + 0 kg	70 kg	55
CAT 4	2600	10.8	60	42,5 kg	5 + 0 kg	5 + 0 kg	85 kg	65
CAT 5	2800	12.2	62	50 kg	5 + 5 kg	5 + 5 kg	100 kg	75

2.15.2 Summary of Research to Date to Underpin PES Development

The former PES was developed in 2010 and consists of static and dynamic exercises. A growing body of literature has questioned the theoretical base for the core test including the connections between core stability, injuries and performance and the reliability and validity of the test methods [86]. The literature supports the practical experience with the core test that in some cases seemed to exclude strong fit soldiers with no injuries while passing weaker unfit soldiers. A work group concluded that the core test had low validity and reliability and that the low diagnostic power causes an unacceptable number of false negatives. Consequently, all static exercises were excluded and a new test was developed.

The aim of the "new" PES was to set a physical minimum requirement that could serve as a basic fitness level for training leading to the operational physical recommendations. There was a requirement for the test to be carried out with as little equipment as possible and easy to conduct with many people and few controllers.

The strength-endurance tests were developed and consisted of four exercises; split squat, dips, horizontal pull-up and a burpee like exercise. The protocol consisted of 2400m run, spilt squat and dips 3x15 reps, horizontal pull-up and burpee 1x15 reps, was used to identify an appropriate level of the test. The protocol was tested on a total of 595 subjects, both men and women in different age groups, fitness levels and job functions. The RPE scale 1-5 was used to measure the intensity of the different exercises. Based on the results, the levels of the different exercises in the PES were determined. (See Section 2.18 for further Background Reading.)

2.16 NORWEGIAN DEFENCE FORCES (ALL BRANCHES)

Approximate date current PES implemented	2017
Approximate duration of the PES	1 – 1.5 hours
Test standards the same for gender	No
Test standards the same for all ages	No
Frequency of PES assessment	Conscripts: at selection (one year prior to service), within the first few weeks of service, and at dismissal after 12 months of service. Officers and professional soldiers: at selection, thereafter once per year
PES user (e.g., tri-service, generic or specialist)	All
Is the PES currently being reviewed or redeveloped	No

2.16.1 Summary of Current PES Processes

The previous Norwegian Defence Forces PES system was developed in the late 1970s, but was replaced by the current system 1st of January 2017. Our current PES system is based on general fitness tests that all military personnel, irrespective of branch, must carry out at selection and thereafter at least once a year. The new regulation also allows inclusion of branch-specific task simulation tests – but few such tests are currently being administered. Aerobic endurance is assessed via a 3000 m run, muscular strength via a standing medicine ball throw (10 kg), standing long jump and pull-ups. Alternative tests used in different occasions are a maximal treadmill walk/run test and the 20 m shuttle run test. Professional soldiers and officers in positions with low physical demands may also select to carry out the annual endurance test using one of three alternative tests: 20 km cycling, 10 km cross-country skiing or 500 m swimming. Each branch sets their own minimum standard levels for their different trades, within a predefined 1 – 9 scale (1 – low physical demands and 9 – high physical demands). The minimum test score requirements are age- and gender-free for trades with minimum requirement 9, but fully adjusted for age and gender for trades with minimum requirement 1. For minimum requirement 8 to 2 the test scores are gradually more age and gender adjusted. The scales have been developed based on a norm-reference system. The branches have been encouraged to carry out their own physical demands analyses prior to deciding their minimum requirements for their trades. Applicants who fail the selection test are not permitted to serve in the military. Soldiers or officers who fail an In-Service test may be relocated to a less physically demanding trade. Each branch may treat subjects who fail differently, according to their specific needs, personnel demands, etc.

2.16.2 Summary of Research to Date to Underpin PES Development

The work related to develop the PES has involved three primary areas:

1) Physical demand analyses;
2) Validity, reliability and practical aspects of various general fitness tests; and
3) Set appropriate minimum requirements for the different trades.

These areas of research were undertaken using:

- A literature review of previous international research.
- Survey to approximately 1000 soldiers and officers pertaining to perceived physical demands of their roles.
- A study of reliability and validity of the muscle strength tests.
- Several smaller validity-studies were also carried out for some of the alternative tests.
- Norm-reference data have been extracted primarily from previous test results, to develop the 1 to 9 scale.

See Section 2.18 for further Background Reading.

2.17 NETHERLANDS ARMED FORCES (NAF) (TRI-SERVICE AND MILITARY POLICE (MP))

Pre-entry physical screening (legal status test)	
Approximate date current PES implemented	2018
Approximate duration of the PES	120 minutes
Test standards the same for gender	Yes
Test standards the same for all ages	Yes
Frequency of PES assessment	Point of selection
PES user (e.g., tri-service, single-service or specialist)	Tri-service and MP
Is the PES currently being reviewed or redeveloped	No
	NAF Physical Fitness Test (legal status test)
Approximate date current PES implemented	2016
Approximate duration of the PES	30 minutes

Test standards the same for gender	No (sex- and age-related standards)
Test standards the same for all ages	No (sex- and age-related standards)
Frequency of PES assessment	Annual
PES user (e.g., tri-service, single-service or specialist)	Tri-service and MP
Is the PES currently being reviewed or redeveloped	Yes
If 'Yes', anticipated date when new PES be implemented	Revision of NAF PF Test is foreseen in 2019

2.17.1 Summary of Current PES Processes

2.17.1.1 Pre-entry (Physical) Screening

All negative scores on the PAR-Q are a pre-requisite for candidates to participate in the physical tests of the newly revised NAF pre-entry screening, comprising running, lifting/carrying, loaded marching, and an indoor test series. Test leaders of the pre-employment centers give instructions on the tests, monitor safe and correct execution of tests, and administer the results. All tests are meant to be individual efforts and have sex- and age-independent 'pass or fail' standards, with job function cluster-dependent standards for the outdoor tests.

These consist of:

- Loaded march, 20 to 48 min with 25 to 45 kg (depending on job cluster).
- Repeated lifting (both floor level and height-adjusted table level) and carrying (25 m) of an ammunition chest with handles, 20/30/40 kg, weights and numbers of repetitions (depending on job cluster).
- Digging sand from one compartment to the other in a bisectional sand bucket (1 m^3) for 1 to 2 min, (depending on job cluster).
- 'Operational endurance' assessment with 12 min run test, with cluster-dependent standards (2200 to 2700 m).
- Working above shoulder level (cluster independent): moving a weight (both hands) from one shaft to another for 1 min at a body height-adjusted height, followed by screwing and unscrewing of a wingnut (left and right hand) for 1 min at the same height.
- Indoor test series (cluster independent):
  - Clambering and scrambling on and off obstacles (no time limits);
  - Squat for 5 seconds;
  - Shooting positions: lying, kneeling left/right;
  - 6 m crawl on hands and knees; and
  - 6 m military crawl.

2.17.1.2 NAF Physical Fitness Test

The commander decides when military members of his/her unit should participate in the annual NAF PF Test for the first time. A passing score has a 365 day currency. Individual military employees are responsible for test participation afterwards.

The testing procedure, supervised by a PTI from the NAF Physical Training and Education Organisation, comprises the following steps:

- Prior to the test, each individual is asked to fill out a checklist 'medical aptitude NAF PFT' (similar to the PAR-Q): if items are scored positively, a physician should be consulted before participating in the test.
- During the test, individuals run through the test modules (unit-wise) in a fixed order: push-ups, sit-ups, and 12-minute run, respectively.
- If the individual fails the test, a 3-month PT program is provided before a next attempt.

The unit commander is responsible for administrating participation rates (yes/no) and test results (pass/fail) into the NAF employee management system. When a military employee has structural medical restraints (i.e., recovery would last longer than 6 months) to take part in the NAF PF Test, the company doctor can advise an alternative health-related fitness test is provided, comprising sit-ups (no alternative), chest-presses (instead of push-ups), and a cardiorespiratory test on a step machine (instead of 12-minute run), with specific sex- and age-related standards. The annual NAF Physical Fitness test is mandatory; not participating may lead to disciplinary consequences. Failing or passing the test has no legal status consequences to date. It is foreseen that the NAF Physical Fitness test will be replaced by a role-related tri-service test in the next coming years. No decision of the NAF senior leadership has been made on this issue thus far.

2.17.2 Summary of Research to Date to Underpin PES Development

The following section outlines the approach taken by the NAF in their PES development over the last two decades. The NAF currently uses PES practices for their pre-employment screening, as well as an annual health-related fitness test with legal status consequences for all services. Other physical testing that are periodically used within each service comprise primarily general or role-related fitness tests for commanders to assess physical readiness, both on an individual level and unit level. These tests are more or less derived from job specific demands but serve as PT-directional tests.

2.17.2.1 Pre-entry (Physical) Screening

In a 2002 Commander-in-Chief Instructions for Regulations [87], 12 Military Basic Requirements (MBR) were identified by the senior leadership, one of them referring to physical capabilities: "A soldier features adequate physical, sensory, mental, and cognitive capabilities to perform in all kinds of operational circumstances, without endangering him/herself and other unit members." In 2009, the **NAF Physical Fitness Test** was installed to annually test minimal physical fitness levels of all NAF military personnel, reflecting this physical MBR. The test comprises a health-related fitness test (12-minute run, push-ups, sit-ups) with gender-specific and age-specific standards, and may have legal status consequences when not participating.

In the first phase of the revision of the **NAF pre-employment screening** (2014 – 2018) [88], an Armed Forces panel of tri-service subject matter experts in expeditionary/domestic deployments translated the MBRs into 26 essential job components, i.e., job characteristics that challenge the physical, mental, and cognitive abilities of the employee, and that have risks for the health and safety of the individual and other unit members which cannot be mitigated with current state-of-the-art measures. The first ten essential job components represented physical requirements, six of which were considered as military critical tasks: loaded marching, clambering and scrambling, lifting loads, carrying loads, digging/working with bended and twisted back, and 'operational endurance'.

Each of these six critical tasks were operationalized into three quasi-operational scenarios with different intensity levels by scientific experts of the Training Medicine and Training Physiology Department of the Royal Netherlands Army. Based on these 3-level critical tasks, six job function clusters were derived (Table 2-11). Job function cluster I basically reflects the 'soldier first' principle.

Table 2-11: NAF Job Function Clusters (Numbers 1-3 Refer to the Intensity Levels).

Job Function Cluster	Loaded Marching	Clambering and Scrambling	Lifting Loads	Carrying Loads	Digging	Operational Endurance
I: Basic soldier	1	1	1	1	1	1
II: Basic soldier with lifting/carrying tasks	1	1	2	2	1	1
III: Basic soldier with heavy lifting/carrying tasks	1	1	3	3	1	1
IV: Soldier with cardiorespiratory challenges	2	2	2	2	2	2
V: Soldier with both heavy lifting/carrying tasks and cardiorespiratory challenges	2	2	3	3	2	2
VI: Special Forces	3	3	3	3	3	3

For the purpose of the pre-entry examination of military applicants, a NAF working group including scientific and medical experts defined so-called medical capacity standards for all 26 essential job components. For the six military critical tasks, different standards were determined per job function cluster. Each medical capacity standard was further operationalized into different instruments for the medical examiners: signalling health questionnaires, biometric measurements, and (mostly) task simulation tests.

The revised NAF pre-entry screening system was implemented in January 2018.

2.17.2.2 NAF Physical Fitness Test

The NAF Physical Fitness Test was first introduced in the Royal Netherlands Army in 1990 as a derivative of the US Army Physical Fitness test, and was implemented as a tri-service NAF test in 2009. Aim of this test is to assess health-related fitness of military NAF employees on an annual base. The test elements represent strength of the upper body ‘stretching chain’ (push-ups), core strength and stability (sit-ups), and general cardiorespiratory fitness (12-minute run). According to the NAF senior leadership, these basic motor skills are essential to perform military critical tasks such as loaded marching, carrying, and sustained readiness. No PES-specific research underlines this assumption.

2.18 BACKGROUND READING

Kimmons, S. (2017). With Six Events, New Army Combat Readiness Test Aims To Replace APFT, Cut Injuries. Army News Service.

Myers, M. (2017). Army Tries Out New Soldier Readiness Test in Pilot at Installations, in Army Times On-line at https://www.armytimes.com/news/your-army/2017/07/21/army-tries-out-new-soldier-readiness-test-in-pilot-at-installations/.

Drain, J., Billing, D., Neesham-Smith, D., and Aisbett, B. (2016). Predicting Physiological Capacity of Human Load Carriage – A Review. Applied Ergonomics, 52, 85-94.

Carstairs, G.L., Ham, D.J., Savage, R.J., Best, S.A., Beck, B., and Billing, D.C. (2018). A Method for Developing Organisation-Wide Manual Handling Based Physical Employment Standards in a Military Context. Journal of Science and Medicine in Sport, 21(11), 1162-1167.

INSTRUCTION N° 126/DEF/EMA/EMP/3 relative au contrôle de la condition physique du militaire. Du 25 janvier 2007.

INSTRUCTION N° 1570/DEF/EMAT/ES/B.EMP/OUT/33 relative au contrôle de la condition physique du militaire pour l'armée de terre. Du 13 novembre 2009.

Hasselstrøm, H. (2017). Rapport vedrørende evaluering af basiskrav

VFKBST FSK 491-7: Bestemmelse for militær fysisk træning og idrætsvirksomhed i FSV

Norwegian Defence University College (2016). Regulations on Physical Fitness Testing. Oslo, Norway.

Kirknes, J., Aandstad, A., and Stornæs, A.V (2014). Innstilling. Revidert fysisk testordning for Forsvaret [Recommendation. Revised Physical Fitness Test System for the Norwegian Armed Forces]. Sognsvann, Oslo, Norway: The Norwegian Defence University College, Norwegian School of Sport Sciences/Defence Institute.

Aandstad, A., and Hageberg, R. (2018). Reliability and Validity of a Maximal Treadmill Test for Predicting Aerobic Fitness in Norwegian Prospective Soldiers. Military Medicine, 184(3-4), e245-e252.

Aandstad. A., and Kirknes. J. (2018). Validity and Reliability of Muscle Strength Field Tests for Use in the Norwegian Armed Forces. 3rd International Conference on Physical Employment Standards. Portsmouth, England.

Aandstad, A. (2014). Predicting Relative and Absolute VO2max from the 3000 meter run. Proceedings of 3rd International Congress on Soldiers' Physical Performance. Boston, USA.

2.19 REFERENCES

[1] Department of National Defence. DAOD 5023-1 (2006). Minimum Operational Standards Related to Universality of Service. Canadian Forces Defence Administrative Orders and Directives. Chief of Military Personnel, Ottawa.

[2] Department of National Defence. Common Military Task Fitness Evaluation (CMTFE) General Description. Canadian Forces Directorate of Fitness (2015). Ottawa.

[3] Department of National Defence. DAOD 5023-0 (2006). Universality of Service. Canadian Forces Defence Administrative Orders and Directives. Chief of Military Personnel, Ottawa.

[4] Department of National Defence. (2016). FORCE Operation Manual: 2nd Edition. Canadian Forces Directorate of Fitness. Ottawa.

[5] Reilly, T., Stockbrugger, B., Saucier, S., Walsh, E., and Gagnon, P. (2016). Physiological Demands of FORCE Combat: A Fitness Objective for the Canadian Army. Proceedings from CIMVHR 2016. Vancouver, BC, Canada.

[6] Department of National Defence. DAOD 5023-2 (2006). Physical Fitness Program. Canadian Forces Defence Administrative Orders and Directives. Chief of Military Personnel, Ottawa.

[7] Department of National Defence. DAOD 5023-0 (2006). Universality of Service. Canadian Forces Defence Administrative Orders and Directives. Chief of Military Personnel, Ottawa.

[8] Blacklock, R.E., Reilly, T.J., Spivock, M., Newton, P.S., and Olinek, S.M. (2015). Standard Establishment Through Scenarios (SETS): A New Technique for Occupational Fitness Standards. Work, 52(2): 375-383. doi:10.3233/WOR-152128. PMID:26409372.

[9] Department of National Defence. (2011). Phase I Report – Identification of Common, Essential, and Physically Demanding Tasks in the CF. Canadian Forces Directorate of Fitness, Ottawa.

[10] Department of National Defence. (2013). Phase II Report – Physical Demands of Common, Essential, and Physically Demanding Tasks in the CF. Canadian Forces Directorate of Fitness, Ottawa.

[11] Constable, S. and Palmer, B. (Eds.) (2000). SOAR: The Process of Physical Fitness Standards Development. Human Systems Information Analysis Centre.

[12] Bonneau, J. (2001). Evaluating Physical Competencies Fitness Related Tests Task Simulation or Hybrid. Objectives, Process and Consensus Summary of The National Forum on Bona Fide Occupational Requirements. In Proceedings of the Consensus Forum on Establishing Bona Fide Requirements for Physically Demanding Occupations. Ed. N. Gledhill, J. Bonneau, and A. Salmon. York University, Toronto, Ont., Canada, pp. 23-33.

[13] Larcom, K., Walker, L.A., Warr, B.J., Smith, L.B., Redmond, J.E., Zambraski, E.J., and Sharp, M.A. (2015). Physical Demands Study-Focus Groups (No. USARIEM-TR-T16-3). US Army Research Institute of Environmental Medicine, Natick, United States.

[14] Boye, M.W., Cohen, B.S., Sharp, M.A., and Canino, M.C. (2016). Responses to Three USARIEM Job Analysis Questionnaires (JAQs) Conducted with Cavalry Scouts and Armor Crewmen (MOSs 19D and 19K) (No. USARIEM-TR-T17-04). US Army Research Institute of Environmental Medicine, Natick United States.

[15] Boye, M.W., Cohen, B.S., Sharp, M.A., and Canino, M.C. (2016). Responses to Three USARIEM Job Analysis Questionnaires (JAQs) Conducted with Cannon Crewmembers and Fire Support Specialists (MOSs 13B and 13F) (No. USARIEM-TR-T17-01). US Army Research Institute of Environmental Medicine, Natick, United States.

[16] Boye, M.W., Cohen, B.S., Sharp, M.A., Hydren, J.R., Simpson, K.S., and Canino, M.C. (2015). Results of Two USARIEM Self-Report Job Analysis Questionnaires (JAQs) Conducted with Combat Engineers (MOS 12B) (No. USARIEM-TR-T16-12). US Army Research Institute of Environmental Medicine, Natick, United States.

[17] Sharp, M.A., Cohen, B.S., Boye, M.W., Foulis, S.A., Redmond, J.E., Larcom, K., ... and Zambraski, E. J. (2017). US Army Physical Demands Study: Identification and Validation of the Physically Demanding Tasks of Combat Arms Occupations. Journal of Science and Medicine in Sport, 20, S62-S67.

[18] Foulis, S.A., Redmond, J.E., Frykman, P.N., Warr, BJ., Zambraski, E.J., and Sharp, M.A. (2017). US Army Physical Demands Study: Reliability of Simulations of Physically Demanding Tasks Performed by Combat Arms Soldiers. The Journal of Strength and Conditioning Research, 31(12), 3245-3252.

[19] Foulis, S.A., Redmond, J.E., Warr, B.J., Sauers, S.E., Walker, L.A., Canino, M.C., ... and Sharp, M.A. (2015). Development of a Physical Employment Testing Battery for Field Artillery Soldiers: 13B Cannon Crewman and 13F Fire Support Specialist (No. USARIEM-TR-T16-9). US Army Research Institute of Environmental Medicine, Natick, United States.

[20] Foulis, S.A., Redmond, J.E., Warr, B.J., Walker, L.A., Canino, M.C., Hydren, J.R., ... and Sauers, S.E. (2015). Development of a Physical Employment Testing Battery for Armor Soldiers: 19D Cavalry Scout and 19K M1 Armor Crewman. US Army Research Institute of Environmental Medicine Natick United States.

[21] Foulis, S.A., Redmond, J.E., Warr, B.J., Zambraski, E.J., Frykman, PN., Gebhardt, D.L., ... and Sharp, M.A. (2015). Development of a Physical Employment Testing Battery for 12B Combat Engineers. Technical Report.

[22] Redmond, J.E., Foulis, S.A., Frykman, P.N., Warr, B.J., Sauers, S.E., Walker, L.A., ... and Sharp, M. A. (2015). Development of a Physical Employment Testing Battery for Infantry Soldiers: 11B Infantryman and 11C Infantryman-Indirect Fire (No. T16-10). US Army Research Institute of Environmental Medicine Natick United States.

[23] Foulis, S.A., Sharp, M.A., Redmond, J.E., Frykman, P.N., Warr, B.J., Gebhardt, D.L., ... and Zambraski, E. J. (2017). US Army Physical Demands Study: Development of the Occupational Physical Assessment Test for Combat Arms Soldiers. Journal of Science and Medicine in Sport, 20, S74-S78.

[24] Sharp, M.A., Foulis, S.A., Redmond, J.E., Canino, M.C., Cohen, B.S., Hauret, K., ... and Gebhardt, D.L. (2018). Longitudinal Validation of the Occupational Physical Assessment Test (OPAT) (No. T18-05). US Army Research Institute of Environmental Medicine Natick United States.

[25] Smith, C.V. (1998). An Ergonomic Analysis of Royal Navy Firefighting. INM Report 98031.

[26] Bilzon, J.L., Scarpello, E.G., Smith, C.V., Ravenhill, N.A., and Rayson, M.P. (2001). Characterization of the Metabolic Demands of Simulated Shipboard Royal Navy Fire-Fighting Tasks. Ergonomics, 44(8), 766-780.

[27] Shvartz, E., and Reibold. R.C. (1990). Aerobic Fitness Norms for Males and Females Aged 6 to 75 Years: A Review. Aviation, Space and Environmental Medicine. 61: 3-11.

[28] Brown, P.E., and Fallowfield, J.L. (2019). Physical Employment Standards for UK Royal Navy Personnel: A Survey of Tasks That Require Muscle Strength and Endurance. Military Medicine. Doi: 10.1093/milmed/usz099.

[29] Brown, P.E.H., and Fallowfield, J. (2013). The Design and Validation of a Strength-Based Royal Navy Fitness Test: Part 2 – A Job Analysis of the Most Critically Demanding Generic Tasks Performed Onboard Royal Navy Ships. INM Report 2013.013.

[30] Brown, P.E.H, Britland, S.E., Stevenson, J., Sloan, B., Simpsin, D., Gunner, F., Fallowfield, J.L. (2013). The Design and Validation of a Strength-Based Royal Navy Fitness Test: Part 3 – Physical Performance Times, Suggested Physical Performance Standards, and Pass/Failure Rates of Five Potential Strength-Based Physical Employment Tests. INM Report 2013.030.

[31] Rayson, M.P., Bell, D.G., Holliman, D.E., Llewellyn, M., Nevola, V.R., and Bell, R.L. (1994). Physical Selection Standards for the British Army Phases 1 and 2. Farnborough, UK: Army Personnel Research Establishment.

[32] Rayson, M.P. and Holliman, D.E. (1994). Physical Selection Standards for the British Army. Phase 3. Development of Physical Selection Tests and Pilot Study. Defence Research Agency.

[33] Rayson, M.P. and Holliman, D.E. (1995). Physical Selection Standards for the British Army. Phase 4. Predictors of Task Performance in Trained Soldiers.: Defence Research Agency.

[34] Rayson, M.P., Holliman, D.E., Nevola, R.V., and Birch, C.L. (1996). Physical Selection Standards for the British Army. Phase 5: Validation.: Defence Research Agency.

[35] Optimal Performance Ltd. (1997). Validation of Physical Selection Standards for Recruits (PSS(R)). Farnham, UK: Optimal Performance Ltd.

[36] Rayson, M., Wilkinson, D., and Nevill, A. (2002). Physical Selection Standards for Single Entry Recruits: Development and Validation Study. Farnham, UK: Optimal Performance Ltd.

[37] Blacker, S.D., Wilkinson, D.M., Brown, P.I., Nevill, A.M., and Rayson, M.P. (2011). Revalidation of Physical Selection Standards for Recruits [PSS(R)] for Junior, Infantry and Soldier Pre-conditioning Courses. Bristol: Optimal Performance Limited.

[38] Blacker, S.D., Wilkinson, D.M., Brown, P.I., and Rayson, M.P. (2011). Revalidation of Physical Selection Standards for Recruits [PSS(R)] Undertaking the Common Military Syllabus (Standard Entry) [CMS(SE)]. Bristol: Optimal Performance Limited.

[39] Blacker, S.D., Wilkinson, D.M., and Rayson, M.P. (2012). Development of Physical Selection Standards for Officer Cadets [PSS(O)]. Bristol: Optimal Performance Ltd.

[40] Blacker, S.D., Wilkinson, D.M., and Rayson, M.P. (2010). Development of Physical Selection Standards for Territorial Army Recruits [PSS(TA)]. Bristol: Optimal Performance Limited.

[41] Tattersall, S., Woolford, K., Pearson, E., and Head, A. (2010). Army Operational Fitness Test Trials: Physiological Data Collection. Porton Down, UK: Dstl.

[42] Blacker, S.D., Myers S.D., Nevola V.R., Rayson M., Walker E., Coakley S., Hale B., Cordell N.J., Knapik J., Gebhardt D.L., Billings A., Mcdevitt C., Monkhouse J. (2016). Interim Technical Report – TIN 3.179: Development and Validation of Role-Related Gender-Free Physical Employment Standards for Ground Close Combat Roles, in the British Ministry of Defence – Phase 1. Defence Human Capability Science and Technology Centre.

[43] Blacker, S.D., Myers S.D., Nevola V.R., Rayson M., Walker E., Coakley S., Gallagher C.A., Lee B.J., Datson N., Doherty J., Flood T.R., Vine C.A.J., Hale B., Cordell N.J., Knapik J., Gebhardt D.L., Nindl B.C., Mcdevitt C., and Monkhouse J. (2017). TIN 3.179: Development and Validation of Role-Related Gender-Free Physical Employment Standards for Ground Close Combat Roles, in the British Ministry of Defence – Phase 2. Defence Human Capability Science and Technology Centre.

[44] Blacker, S.D., Myers S.D., Walker E., Coakley S., Gallagher C.A., Lee B.J., Doherty J., Flood T.R., Vine C.A.J., Rayson M.P., Hale B., Cordell N.J., Knapik J., Gebhardt D.L., Nindl B.C., Mcdevitt C., and Monkhouse J. (2018). TIN 3.179: Development and Validation of Role-Related Gender-Free Physical Employment Standards for Ground Close Combat Roles, in the British Ministry of Defence – Phase 3. Defence Human Capability Science and Technology Centre.

[45] Blacker, S.D., Myers S.D., Walker E.F., Lee B.J., Rue C.A., Coakley S.L., Needham-Beck S., Siddall A., Doherty J., Flood T.R., Vine C.A.J., Rayson M., Mcdevitt C., and Hale B. (2018). TIN 3.179 Development and Validation of Role-Related Gender Free Physical Employment Standards (PES) for

Ground Close Combat (GCC) Roles in the British Ministry of Defence (MOD) – Phase 3c – Impact Analysis and Additional Non-GCC Women's Trial Final Report. Defence Human Capability Science and Technology Centre.

[46] Myers, S.D., Blacker S.D., Coakley S., Walker E., Gallagher C.A., Lee B.J., Doherty J., Flood T.R., Vine C.A.J., Rayson M.P., Hale B., Cordell N.J., Knapik J., Gebhardt D.L., Nindl B.C., Mcdevitt C., and Monkhouse J. (2019). TIN 3.179: Development and Validation of Role-Related Gender-Free Physical Employment Standards for Ground Close Combat Roles, in the British Ministry of Defence – MJP4 Report. Defence Human Capability Science and Technology Centre.

[47] Blacker, S.D., Myers S.D., Coakley S., Rue C.A., Walker E., Lee B.J., Siddall A.G., Needham-Beck S.C., Doherty J., Flood T.R., Vine C.A.J., Ashdown K.M., Hinde K.L., Osofa J.I., Walker F.S., Rayson M.P., Hale B., Cordell N.J., Knapik J., Gebhardt D.L., Nindl B.C., McDevitt C., and Monkhouse J. (2019). TIN 3.179: Development and Validation of Role-Related Gender-Free Physical Employment Standards for Ground Close Combat Roles, in the British Ministry of Defence – Phase 4. Defence Human Capability Science and Technology Centre.

[48] Blair, N., Kohl, H., Paffenbarger, R., Clark, D., Cooper, K., and Gibbons, L. (1989). Physical Fitness and All- Cause Mortality. A Prospective Study of Healthy Men and Women. JAMA. 262(17).

[49] Kasch, F.W., Boyer, J.L., Van Camp, S.P., Verity, L.S., and Wallace J.P. (1993). Effect of Exercise on Cardiovascular Ageing. Age and Ageing, 22 (1:5-10).

[50] Nindl, B.C., Jones, B.H., Van Arsdale, S.J., Kelly, K., and Kraemer W.J. (2016). Operational Physical Performance and Fitness in Military Women: Physiological, Musculoskeletal Injury, and Optimized Physical Training Considerations for Successfully Integrating Women into Combat-Centric Military Occupations. Military Medicine, 181 (1:50).

[51] Drain, J.R., Aisbett, B., Lewis, M., and Billing, D.C. (2017). The Pandolf Equation Under-Predicts the Metabolic Rate of Contemporary Military Load Carriage. Journal of Science and Medicine in Sport, 20, S104-S108.

[52] Tofari, P.J., Treloar, A.K.L., and Silk, A.J. (2013). A Quantification of the Physiological Demands of the Army Emergency Responder in the Australian Army. Military Medicine, 178(5), 487-494.

[53] Treloar, A.K.L., and Billing, D.C. (2011). Effect of Load Carriage on Performance of an Explosive, Anaerobic Military Task. Military Medicine, 176(9), 1027-1031.

[54] Silk, A.J., and Billing, D.C. (2013). Development of a Valid Simulation Assessment for a Military Dismounted Assault Task. Military Medicine, 178(3), 315-320.

[55] Taylor, N.A., Burdon, C.A., Van Den Heuvel, A.M., Fogarty, A.L., Notley, S.R., Hunt, A.P., ... and Peoples, G.E. (2015). Balancing ballistic Protection against Physiological Strain: Evidence from Laboratory and Field trials. Applied Physiology, Nutrition, and Metabolism, 41(2), 117-124.

[56] Beck, B., Middleton, K.J., Carstairs, G.L., Billing, D.C., and Caldwell, J.N. (2016). Predicting Stretcher Carriage: Investigating Variations in Bilateral Carry Tests. Applied Ergonomics, 55, 124-132.

[57] Beck, B., Carstairs, G.L., Billing, D.C., Caldwell, J.N., and Middleton, K.J. (2017). Modifiable Anthropometric Characteristics are Associated with Unilateral and Bilateral Carry Performance. The Journal of Strength and Conditioning Research, 31(2), 489-494.

[58] Beck, B., Ham, D.J., Best, S.A., Carstairs, G.L., Savage, R.J., Straney, L., and Caldwell, J.N. (2016). Predicting Endurance Time in a Repetitive Lift and Carry Task Using Linear Mixed Models. Plos One, 11(7), E0158418.

[59] Beck, B., Carstairs, G.L., Caldwell Odgers, J.N., Doyle, T.L., and Middleton, K.J. (2016). Jerry Can Carriage is an Effective Predictor of Stretcher Carry Performance. Ergonomics, 59(6), 813-820.

[60] Carstairs, G.L., Ham, D.J., Savage, R.J., Best, S.A., Beck, B., and Doyle, T.L. (2016). A Box Lift and Place Assessment is Related to Performance of Several Military Manual Handling Tasks. Military Medicine, 181(3), 258-264.

[61] Savage, R.J., Best, S.A., Carstairs, G.L., and Ham, D.J. (2012). The Relationship Between Maximal Lifting Capacity and Maximum Acceptable Lift in Strength-Based Soldiering Tasks. The Journal of Strength and Conditioning Research, 26, S23-S29.

[62] Savage, R.J., Best, S.A., Carstairs, G.L., Ham, D.J., and Doyle, T.L. (2014). On the Relationship Between Discrete and Repetitive Lifting Performance in Military Tasks. The Journal of Strength and Conditioning Research, 28(3), 767-773.

[63] Beck, B., Middleton, K.J., Billing, D.C., Caldwell, J.N., and Carstairs, G.L. (2019). Understanding Anthropometric Characteristics Associated with Performance in Manual Lifting Tasks. The Journal of Strength and Conditioning Research, 33(3), 755-761.

[64] Middleton, K.J., Carstairs, G.L., and HAM, D.J. (2016). Lift Performance and Lumbar Loading in Standing and Seated Lifts. Ergonomics, 59(9), 1242-1250.

[65] Middleton, K.J., Carstairs, G.L., Caldwell, J.N., Billing, D.C., and Beck, B. (2017). The Sensitivity of a Military-Based Occupational Fitness Test of Muscular Strength. Applied Ergonomics, 60, 255-259.

[66] Kilding, H. and Fordy, G. (2014). Development of a Land Combat Fitness Test for the New Zealand Army. DTA Report 388.

[67] Dobson, B. (2012). Land Combat Fitness Test Fire and Manoeuvre Trials. DTA File No. 3750/A0682/20.

[68] Kilding, H. and Fordy, G. (2014). Test-retest Reliability Assessment of the Land Combat Fitness Test. DTA Minute 3750/J1380/26March2014.

[69] Mcara, P.G. (2003). RNZAF Fitness Testing. Report Mar 13, 2003.

[70] Tourell, B.M. (2006). Physical Fitness in the RNZAF. RNZAF Minute, May 11 2016. RNZAF, Ohakea.

[71] Kilding, H. (2017). Fitness Testing in the New Zealand Defence Force: A 2017 Review. DTA Technical Note 2017/1.

[72] Kilding, H. (2015). Development of a Valid and Reliable Operational Fitness Test for the Royal New Zealand Navy. DTA Technical report 2015/1.

[73] Kilding, H. (2012). Estimation of the Physical Demand of the Royal New Zealand First Aid/EEBD run. DTA Report 344.

[74] Kodama, S., Saito, K., Tanaka, S., Maki, M., Yachi, Y., Asumi, M., ... and Yamada, N. (2009). Cardiorespiratory Fitness as a Quantitative Predictor of All-Cause Mortality and Cardiovascular Events in Healthy Men and Women: a Meta-Analysis. JAMA, 301(19), 2024-2035.

[75] Reilly, T., Spivock, M., and Prayal-Brown, A.L. (2014). The Fitness Profile: A Case for Assessing Aerobic Fitness and Body Composition within the FORCE Evaluation. Technical Report: Department of National Defence, Assistant Deputy Minister (Science and Technology). Ottawa.

[76] Leyk, D., Witzki, A., Gorges, W., Erley, O., Schlattmann, A., et al. The Basis-Fitness-Test (BFT) and its Data-Management System: A new tool to Monitor Physical Fitness and to Analyse Performance Predictors. Presented at: 2nd International Congress on Soldiers' Physical Performance. Jyväskylä: University of Jyväskylä, 2011.

[77] Rüther, T., Leyk, D., Kahl, L.M., and Dobmeier, H. Exercise-physiological evaluation of the BFT Rating System and Establishing Means of Quality Control. Technical Report 2012 Annual Military Scientific Research Report, Defence Research for the German Armed Forces, Bonn, Germany, 2013.

[78] Rüther, T., Sievert, A., Trunzler, M., Gorges, W., and Leyk, D. Exercise-Physiological Evaluation of the BFT Rating System and Establishing Means of Quality Control. Technical Report, Defence Research for the German Armed Forces, Bonn, Germany, 2014.

[79] Eßfeld, D., Rüther, T., Wunderlich, M., and Sievert, A. Development of Achievement Tests and Test Procedure Based Upon Deployment Conditions. Technical Report, Defence Research for the German Armed Forces, Bonn, Germany, 2007.

[80] Leyk, D., Rohde, U., Gorges, W., Rüther, T., and Witzki, A: Physical Fitness of German soldiers 2010 – 2015. J Sci Med Sport, 20: S11, 2017.

[81] Rohde, U., Rüther, T., and Leyk, D. Basic Military Fitness Tool (BMFT): A Reliable Field Uniform-Test for Performance Prediction of Strength-Related Common Military Tasks. J Sci Med Sport, 20: S170, 2017.

[82] Rohde, U., Rüther, T., Hackfort, D., and Leyk, D. Basic Military Fitness Tool (BMFT): Technical Report 2014 Annual Military Scientific Research Report, Defence Research for the German Armed Forces, Bonn, Germany, 2015.

[83] Rohde, U., Sievert, A., Rüther, T., Witzki, A., and Leyk, D. Concept for a Predeployment Assessment of Basic Military Fitness in the German Armed Forces. J Strength Cond Res, 29: S211-S215, 2015.

[84] Rohde, U., Rüther, T., and Leyk, D: Basic Military Fitness: Validation of a Pre-Deployment Assessment Tool. J Sci Med Sport, 20: S78-S79, 2017.

[85] Sievert, A. Fitness Testing in the Bundeswehr – Tracking Individual and Force-Wide Baseline Fitness Using Quality-Controlled BFT Data. Presented at: 3rd International Congress on Soldiers' Physical Performance, Boston, USA, 2014.

[86] Wirth, K., Hartmann, H., Mickel, C., Szilvas, E., Keiner, M., and Sander, A. (2017). Core Stability in Athletes: A Critical Analysis of Current Guidelines. Sports Medicine, 47(3), 401-414.

[87] Ministeriële Regeling A-700a: Militaire Basiseisen (Den Haag, 2002).

[88] Aanwijzing Militaire Gezondheidszorg: De Militaire Aanstellingskeuring (Utrecht, 2017).

Chapter 3 – NATO GUIDE FOR PES DEVELOPMENT

T.J. Reilly
Department of National Defence
CANADA

S. Blacker
University of Chichester
UNITED KINGDOM

M. Sharp
US Army Research Institute for Environmental Medicine
UNITED STATES

D. Gebhardt
Human PRO
UNITED STATES

P. Brown
UK Ministry of Defence
UNITED KINGDOM

J. Drain
Defence Science and Technology Group
AUSTRALIA

H. Kilding
New Zealand Defence Force
NEW ZEALAND

3.1 INTRODUCTION

There are multiple publications which describe the process for developing an evidence-based, role-related, scientifically defensible Physical Employment Standard (PES) in a series of steps or phases [1], [2], [3], [4], [5], [6], [7]. The procedures for developing PES specifically for military occupations in the UK, Canada, US, and Australia have been outlined and discussed by Reilly et al. [6], and have been updated and modified by this working group as depicted in Figure 3-1. This figure is a graphical depiction of the contents of this chapter. Each step of the process will be discussed in detail with examples and case studies to illustrate their importance in military PES.

3.2 UNDERSTANDING THE ENVIRONMENT

3.2.1 Establishment of the Project Management Team / Military Judgment Panel

Before beginning the research process associated with PES development, it is important to establish a Project Management Team (PMT) and/or Military Judgment Panel (MJP) consisting of representatives from the research team and the employer. These include representatives from employment equity, mission planners, labor law representative (Legal counsel), and military personnel involved in selection, training and employment category management. Throughout the PES development process the PMT/MJP should assist in facilitating the recruitment and participation of incumbent workers and potential applicants as research subjects. Other members of the PMT should consist of important stakeholders from the organisation, as well as representatives from those who will be delivering the PES (e.g., military physical training instructors or civilian exercise physiologists), and military health/medical representatives.

3.2.2 PES Governance Case Studies

It is important that the PMT/MJP understand the role of PES within their national legislation. For example, PES development is governed by a Charter of Human Rights (CANADA) [8], Equality Act (UNITED KINGDOM) [9], Civil rights employment law (UNITED STATES) [10], the Equality and Anti-Discrimination Act (NORWAY) [11], the Discrimination Act of 1992 (AUSTRALIA) [12], and The Equal Treatment Act (1994) *[Algemene wet gelijke behandeling, 1994] in The Netherlands* [13]. The aim of this legislation is to ensure that employment standards are fair and reasonable and based on a bona fide occupational requirement. It is very important that this is conveyed to all members involved in the research process, especially the PMT/MJP.

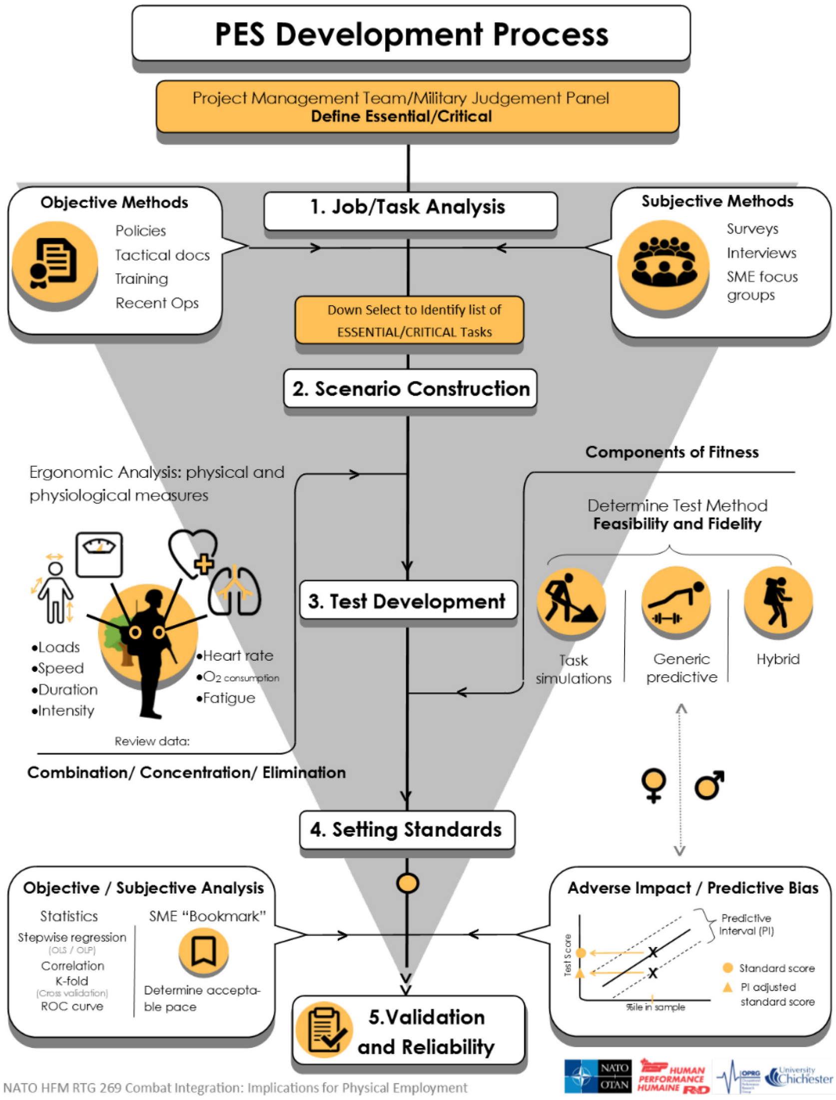


Figure 3-1: PES Development Process.

These laws and acts are very similar in nature, for example France's constitution prohibits discrimination based on sex, race, belief, and trade union activity. Likewise, the Australian laws (e.g., The Discrimination Act of 1992) [12] address the same discrimination factors (e.g., sex, age, religion). The Norwegian Defence University College has the authority to develop and control the test system used in the Norwegian Armed Forces, and military lawyers review the regulations prior to publication and implementation. Yet, the test system still has to be developed along the lines of the civilian Equality and Anti-Discrimination Act.

This section provides more detail about the relevant legislation related to PES in the UK, Canada, and the US, while giving specific examples of how this legislation has been tested in the courts. In general, when determining what will and won't be defensible it is prudent to review examples from case law, such as the Case Studies included here.

3.2.2.1 United Kingdom

Physical Employment Standards in the UK are subject to a number of legislative requirements as outlined in the Equality Act 2010. The Equality Act covers nine key protected characteristics (age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex and sexual orientation). In the context of employment, this equality legislation requires that there is no unlawful discrimination because of a protected characteristic (e.g., a person's sex or age). Any test that is used to select or retain individuals for a job comes under the requirements of the employment provisions of the Equality Act 2010.

The nature of a PES means that there is potential to discriminate unlawfully both directly or indirectly. Direct discrimination occurs if someone is treated less favorably because of a protected characteristic. Indirect discrimination occurs if a provision, criterion or practice (such as a PES) puts someone from a protected group at a disadvantage and is unable to be justified as being a proportionate (fair and reasonable) means of achieving a legitimate aim (e.g., maintaining operational effectiveness). Only an Employment Tribunal can decide on what constitutes "a proportionate means" and "a legitimate aim". The arguments for having different physical selection standards for men and women in the past usually related to the concepts of sex fairness and/or health-related fitness, not occupational standards. However, setting a lower fitness standard for women than for men would result in unlawful direct discrimination against men who are denied a job if they passed the women's pass mark, but not the men's. Direct sex discrimination cannot be justified in law. An example of this is described in the Case Study (Allcock v The Chief Constable of Hampshire Constabulary – 1997). See Figure 3-2.

PES CASE LAW: UK

Allcock v The Chief Constable of Hampshire Constabulary (1997)

Allcock was a serving Police Officer who applied to join the dog section, to do so he had to complete a 2-mile multi-terrain course which required a gender-fair pass standard of 17:00 min:s for women and 16:00 min:s for men. Allcock achieved a time of 16:45 min:s (i.e. failed the standard for men but passed the standard for women), he challenged the standard on the basis that the demands of the role of an Officer in the dog section were the same for men and women and thus the test standards should also be equal for each sex. The employment tribunal concluded, "In failing to conduct a gender neutral test to establish whether a particular candidate is capable of undertaking the duties of a dog handler, the respondents have unfairly directly discriminated against the applicant on the grounds of his sex."

Figure 3-2: PES Case Law: UK – Allcock v The Chief Constable of Hampshire Constabulary (1997).

Hence, to comply with the Equality Act, any PES must reflect the essential physical tasks required to perform the specified job successfully. The pass standards on PES must also reflect the minimum physical performance standards required to complete these essential/critical job tasks. With regard to indirect discrimination, there have been two cases where PES has been challenged on these grounds, these are described in the Case Studies included here from the UK of Bamber v Greater Manchester Police (2010) and Dougan v Chief Constable RUC (1997/98), Figure 3-3. Both of these cases were determined by the employment tribunal as discriminatory.

PES CASE LAW: UK

Bamber v Greater Manchester Police (2010)

In the case of Bamber v Greater Manchester Police (2010) a female Police Inspector failed to meet the gender-free pass standard of 2:45 min:s on a 500 m shield run test and claimed the test showed no resemblance to the operational policing role and she had been put at a disadvantage on test due to sex and age. The tribunal accepted that the claimant had been disadvantaged due to sex and age but that the Greater Manchester Police have a legitimate aim of ensuring appropriate levels of fitness to ensure safety in public order situations. The tribunal accepted that it was legitimate to have a single fitness standard for men and women, provided that the standard is sufficiently related to L2 public order role. Although the content of the test was seen to be sufficiently related to the role there was no evidence for the 2:45 min:s pass standard so the tribunal were not satisfied that the test demonstrated a proportionate means of achieving legitimate aim and therefore the claim for sex and age discrimination was well founded.

Dougan v Chief Constable RUC (1997/98)

In the second case of a PES being challenged on the grounds of indirect discrimination, a female Officer in the Royal Ulster Constabulary (RUC) Reserve applied to join the Regular RUC, and had to pass a series of physical tests as part of the selection process (Dougan v Chief Constable RUC (1997/98)). The claimant failed to meet the gender free pass mark of 3:45 min:s for the Physical Competence Assessment (PCA) which involved a circuit including movements and actions to replicate the physical requirements of an Officers role. A claim of unlawful indirect discrimination was made on the basis that the PCA was not justifiable. The tribunal concluded that the job related fitness test was justifiable for Police Officers, and approved the way in which the test had been developed by a working party and that the PCA properly reflected core activities and competencies of job. However, the tribunal deemed that the PCA pass time did not reflect performance of existing women Officers from a validation exercise and hence was not justifiable, resulting in unlawful indirect sex discrimination. There was also debate in the case regarding how the pass time had been set using performance data and the best judgment of subject matter experts.

Figure 3-3: PES Case Law: UK – Bamber v Greater Manchester Police (2010).

3.2.2.2 Canada

Under the Bona fide Occupational Standards for Employment (Government of Canada, 1985) Human Rights legislation is established to ensure that all individuals have equal opportunity to employment without being hindered by discriminatory practices. However, a discriminatory test such as a PES can be defensible as “it is not a discriminatory practice if any refusal, exclusion, expulsion, suspension, limitation, specification or preferences in relation to employment by an employer is based on a Bona fide Occupational Requirement (BFOR) (Canadian Charter of Human Rights, 1985; Government of Canada, 1985)". The basis for what justifies a PES as a BFOR was determined as a result of the Meiorin Test (1999) and was applied in the context of the DND Fire Fighters in 2006, described here in Case Study, Barr v Treasury Board 2006 (Figure 3-4).

PES CASE LAW: CANADA

Barr vs Treasury Board
(Dept National Defence) 2006

In 2006 the Canadian Armed Forces (DND) required its firefighters, as a condition of continued employment, to complete a fitness test (a task-based circuit consisting of 10 tasks) in 8 minutes or less. This standard was grieved by incumbent firefighters T Barr and S Flannery, who had not attempted the test but contented that a failure to meet this standard could result in carrier action including counselling, demotion and/or termination. Ms Flannery explained to the court that actual fire fighting and rescue operations, as shown in her job description, corresponded to only 4% of her duties.

When this fitness test was developed in 1994 data were collected from 202 male and 7 female DND firefighters, as well as 17 municipal firefighters to increase female participation. The 8 minute standard was justified as a performance objective which would provide a challenge for those young aerobically fit members, while being an obtainable goal for older less fit members. At the 8 minute standard the average VO_{2max} of the failing group was 39 ml/kg/min and 44 ml/kg/min for the passing group. The authors indicated that a VO_{2max} of between 39-45 ml/kg/min was recommended in the literature for firefighters.

Throughout the legal challenge an external researcher analysed the raw data from the 1994 study and deduced that a predicted VO_{2max} of 44ml was only 50 % accurate in identifying those who could and could not attain the 8 minute standard. In addition he indicated that at the 8 minute standard 78% of women and 100% of men over 50 would fail. A training study was conducted, which determined that for 9 women with training, circuit time was reduced from 8:52 to 6:56 (min:sec).

However, this trial concluded that the 8 minute standard was not a BFOR and was prima facia a discriminatory standard on the basis of age and sex, as it does it was chosen arbitrarily and there was no evidence to support the view that the 8 minute standard is the minimum to perform the tasks of a firefighter safely and efficiently.

Figure 3-4: PES Case Law: Canada – Barr vs Treasury Board (Department of National Defence) 2006.

The Canadian Armed forces operates as a unified institution consisting of sea, land, and air elements referred to as the Royal Canadian Navy (RCN) https://en.wikipedia.org/wiki/Royal_Canadian_Navy, Canadian Army (CA) https://en.wikipedia.org/wiki/Canadian_Army, and Royal Canadian Air Force (RCAF) https://en.wikipedia.org/wiki/Royal_Canadian_Air_Force. To ensure members are capable of deploying on joint operations and performing tasks at a minimum performance standard the Canadian Forces (CAF) principle of Universality of Service holds that all personnel must be capable at all times of performing a broad range of general military tasks, common defence and security duties, in addition to the specific duties associated with their occupations.

Universality of Service has a legal basis. It is imposed by Section 33(1) of the *National Defence Act*, which states that all Regular Force members are "at all times liable to perform any lawful duty." If this criteria is not met they lose the right to serve in the Regular Force except during a carefully limited period of recovery from injury or illness as a period of transition out of the military and into civilian life. The Universality of Service principle is also known as the "soldier first" principle, identifying the men and women of the CAF as members of the profession of arms first. Before they are identified as pharmacists, logistics officers or pilots every member, regardless of their military occupation, or whether their place of work is a desk, a ship or the cockpit of an aircraft, must meet the Universality of Service standards in order to remain in the CAF. The "solider first" principle is not specific to Canada, the meaning of this principle across nations was discussed and debated at the 2018 International Conference for Physical Employment Standards, and a review discussing "soldier first" is included in Chapter 8 as a result.

Previous case law provides us direction in that a task does not have to be likely in order for the CAF to include it in the Universality of Service standards. In Case Study Jones v Canada (Attorney General, 2009) "the low likelihood of being deployed cannot trump the Universality of Service Principle" and "it did not matter whether Mr. Jones was likely or not to be deployed, and if so, where he would be posted" (Figure 3-5). Therefore, tasks common to all CAF personnel may include those that may or may not be required to complete in the daily work requirements, during deployment, or during special operations. In summary, the PES for the CAF is applicable to all personnel regardless of sex, gender, age or occupation, and represents the minimum level of fitness required for service in the CAF based on tri-service or the Universality of Service.

PES CASE LAW: CANADA

Jones v Canada (2009)

Mr Jones served as a marine engineer in the Canadian Armed Forces for almost 30 years. He brought an application for judicial review to challenge a decision to release him from the Forces on medical grounds.

Mr. Jones' argument was that his limitations could not impact him as he was not likely to be sent to sea , this was rejected. The court dismissed Mr. Jones' argument that his deployability need only be assessed once the decision to post him in a particular assignment has been made. The low likelihood of being deployed cannot trump the Universality of Service principle. This principle is firmly embedded in section 33 of the *National Defence Act*.

The fact that Mr. Jones may have performed well and met all the requirements of his job was not an indication that he could be posted somewhere else and that he would encounter no problems despite his medical employment limitations.. Therefore, the conclusion was that he was in breach of the Universality of Service principle, and the decision to release him was determined as "not unreasonable".

The fundamental importance of this principle to the functioning and effectiveness of the CAF is recognized in subsection 15(9) of the *Canadian Human Rights Act*.

Figure 3-5: PES Case Law: Canada – Jones v Canada (2009).

3.2.2.3 United States

A recent memorandum explains that Title VII of the Civil Rights Act of 1964, which was enacted to prohibit employment practices that discriminated on the basis of "race, color, religion, sex, or national origin." applies to non-federal employers (42 U.S.C. § 2000e-2(a)) and to federal agencies (42 U.S. Code § 2000e-16(a)). However, Title VII does not apply to uniformed military personnel. Although 42 U.S.C. § 2000e-2(a) specifically includes military departments under Title V, US Code, uniformed military personnel are governed by Title X, US Code. The Code of Federal Regulations section that governs the Equal Employment Opportunity Commission's handling of Title VII **specifically excludes members of the armed forces** (29 C.F.R. § 1614.103(d)(1)).

In addition the United States holds an exemption from the 29 C.F.R. § 1630.2(n)(3) if there is risk of a Direct Threat. A Direct Threat means a significant risk of substantial harm to the health or safety of the individual or others that cannot be eliminated or reduced by **reasonable accommodation**. The determination that an individual poses a "direct threat" shall be based on an individualized assessment of the individual's present ability to safely perform the **essential functions** of the job. An interesting Case Study is presented in Figure 3-6 on the topic of developing PES for the Police in the United States. While these regulations are not always applicable to the military as explained above, this Case Study is a good illustration of the complexity of defining a defensible minimum standard for a public safety occupation.

Overall the UK, Canada and the US have defined in the law what is considered a bona fide or genuine occupational requirement. Many PES, even military ones, must meet these requirements to be considered legally valid in order to discriminate (Case Study: BFOR/BFOQ/GOQ). See Figure 3-7.

The difference in the legalities of civilian vs military PES lie in the requirement of the employer to accommodate. In this way, both the US and Canada have regulations which allow them to discriminate based on the nature of military service be it the "Universality of Service" principle (or Soldier First principle), in Canada or the "Direct Threat" exemption. Without these exemptions the employer (the military) may be responsible for accommodating the employee should they be unable to physically perform the critical/essential job requirements.

3.3 JOB/TASK ANALYSIS

The job/task analysis includes task identification and task qualification, those tasks which are defined as Essential/Critical will qualify for the foundation of a PES.

3.3.1 Defining an Essential/Critical Task

As described by legislation around the world and the case studies provided, for a PES to be scientifically valid and legally defensible, they must accurately represent essential/critical tasks which are physically demanding requirements within the job [6], [7], [14], [15]. To achieve this, the employer must establish and agree on a definition of *Essential/Critical*. *Essential* is more typical terminology in North American literature and *Critical* in European Literature [5], [16]. Through the initial stages of the PES research, the researcher and employer will work with job specific SME to determine which tasks qualify as *Essential/Critical* and why, while referencing this definition as agreed upon in the PMT. Therefore, once the PMT is established their first task should be to agree upon a definition of *Essential/Critical* which fits the specific work environment, perhaps for which the failure to perform to a minimum standard would result in for example, risk to life or safety and/or failure to accomplish the task. Case Study: Definition of Essential provides examples of established definitions of Essential and Critical for Canada, the United States, and the United Kingdom. See Figure 3-8.

Often emergency and lifesaving requirements of the job will be the basis for the PES as they are difficult to dispute and the ability to perform these without technology is accepted so as to ensure they can be carried out successfully under all circumstances [17].

Once the PMT has established a definition of *Essential/Critical* it should be reviewed by legal representatives and formally endorsed and documented by the PMT/MJP. This endorsement is highly recommended as it provides the reference point for the proceeding steps in PES development. Furthermore, without this written acknowledgement, obtaining acceptance of the final product from military members may be problematic.

PES CASE LAW: USA

Catherine Lanning (the plaintiff) vs SEPTA (Southeastern Pennsylvania Transportation Authority) (1989)

In 1989 it was discovered that SEPTA Transit Police Department was unable to control crime on SEPTA property and that problems existed with the physical fitness and capabilities of its transit police officers. At this time there were no PES or physical training programs in place for SEPTA officers. As a result, there were instances where officers were injured, and there were numerous cases of police brutality that were caused by officers retaliating against persons who had previously assaulted physically unfit police officers.

Dr. Davis consulted Subject Matter Experts of SEPTA who described the tasks involved in performing their duties as SEPTA transit officers. The SME then determined the relative criticality/importance/frequency of the tasks on a scale from one to five or six. Throughout these discussions the SME stated that it was reasonable to expect SEPTA officers to have to run one mile in full gear in 11.78 minutes. Dr. Davis, however, rejected this information when creating a 1.5 mile run as a component of SEPTA's physical fitness test, feeling that this pace was too slow as it would require an aerobic capacity that almost any person could meet.
Therefore, Dr. Davis suggested that SEPTA implement a distance running test whereby applicants would be required to run 1.5 miles in 12 minutes or less, requiring 42.5 mL/kg/min which Dr Davis had also recommended for the Anne Arundel County Police.

From 1991-1996 this 42 mL/kg/min was hotly debated by multiple experts in court. Once longitudinal data existed on SEPTA officer performance and their performance on the PES an expert, Dr. Siskin, testified and showed that when comparing officers who were always at 42 mL/kg/min or over to officers who were always under 42 mL/kg/min, the higher aerobic capacity group had a 57.1% "arrest rate" advantage in the more serious crimes and 28% greater arrest rate for all offenses. In addition they made three times (151%) the actual number of Part I arrests and 75% more actual overall arrests when compared to officers who never met the 42 mL/kg/min requirement.

For the years 1991, 1993, and 1996, an average of only 12% of women applicants passed SEPTA's 1.5 mile run in comparison to the almost 60% of male applicants who passed. On January 28, 1997, after satisfying all administrative prerequisites, five women who failed SEPTA's 1.5 mile run brought a Title VII class action against SEPTA on behalf of all 1993 female applicants, 1996 female applicants and future female applicants for employment as SEPTA police officers who have been or will be denied employment by reason of their inability to meet the physical entrance requirement of running 1.5 miles in 12 minutes or less. The links between SEPTA officer performance on the PES of 42mL/kg/min were made through a costly process of multiple professional exercise physiologists and expert testimony applying various statistical analyses post-hoc on incumbent performance and their PES score. This costly procedure could have been avoided if the original standard was based on a physiological analysis of the Essential tasks and at the minimum standard before PES implementation.

Figure 3-6: PES Case Law: USA – Catherine Lenning (the Plaintiff) vs SEPTA (Southeastern Pennsylvania Transportation Authority) (1989).

CASE STUDY:
BFOR/BFOQ/GOQ

In Canada there is a Bona Fide Occupational Requirement (BFOR), in the US a Bona Fide Occupational Qualification (BFOQ), and in the UK a Genuine Occupational Qualification (GOQ). For these three countries, this legislation allows the employer to consider making decisions about the hiring and retention of employees requiring specific physical abilities, where this may otherwise be considered discrimination. Having employment equity advice through the PES development process may help to prevent legal challenges in the future and will ensure these groups understand that if discrimination does occur, it is justified based on the BFOR/BFOQ/GOQ principle.

Figure 3-7: Case Study: BFOR/BFOQ/GOQ.

CASE STUDY:
DEFINITION OF ESSENTIAL
Canada/US/UK

In the Canadian Armed Forces (CAF) the definition of Essential was developed by the PMT, with direction and assistance from the legal team. The definition was determined as a task for which failure would have the potential to:

(1) Cause Injury or death to the military member or member of the general public

(2) Compromise the outcome of a mission/operation

(3) Cause significant damage to crown property

Reilly et al (2013)

This definition was adopted by the UK in 2016 (Blacker et al, 2016)

Similarly, For the purposes of Industry, US law stipulates that PES be based on an "Essential Function" of the job. Essential functions/tasks are defined in 29 C.F.R. § 1630.2(n)(1) as those that are "fundamental" not "marginal." The regulations further set forth a non-exhaustive list of seven examples of evidence that are designed to assist a court in identifying the "essential functions" of a job. They include (1) The employer's judgment as to which functions are essential; (2) Written job descriptions prepared before advertising or interviewing applicants for the job; (3) The amount of time spent on the job performing the function; (4) The consequences of not requiring the incumbent to perform the function; (5) The terms of a collective bargaining agreement; (6) The work experience of past incumbents in the jobs; and/or (7) The current work experience of incumbents in similar jobs. 29 C.F.R. § 1630.2(n)(3). A civilian example of this application is described in a long standing case Lanning v SEPTA (1989-1996) where multiple resources were used to justify an aerobic standard for the police force (see case study Lanning v SEPTA, 1989).

Figure 3-8: Case Study: Definition of Essential Canada/US/UK.

These *Essential/Critical* tasks are often also called Criterion tasks. A Criterion Task being a physically demanding task that is considered to be essential for safe and effective performance in a specified job or role. Definitions such as these compel researchers and employers to ensure that the standards are criterion-based and validly linked to the *Essential/Critical* physically demanding tasks of the occupation. For those in front line combat military trades, job descriptions and obligations are often well defined, however, for support trades, unless there is a "soldier first" principle, these essential/critical tasks are less simple to relate to front line duties (Section 33(2) of the National Defence Act, 1985).

In summary, the open-ended nature of military service is one of the features that distinguish it from the civilian notion of employment governed by a contract, which obliges employees to perform only those duties specified in their job description or contract. Those tasks linked to safety will always be more easily defensible than those linked to cost, as the damage or loss in any given case must be so significant that it would constitute undue hardship to the employer before it would provide the employer with an exception to their duty to accommodate.

3.3.2 Down Selection

Once the PMT/MJP is established, the national legislation identified and a definition of *Essential/Critical* determined, the process of understanding the job can begin. The **PES Development Process Illustration** describes the beginning as a more subjective understanding of the complete "job" using experienced opinions and concludes with an objective representation of the job based on primary data collection and observed practice [4], [5], [6]. This filtration process is also illustrated in Figure 3-9 taken from Ref. [6].

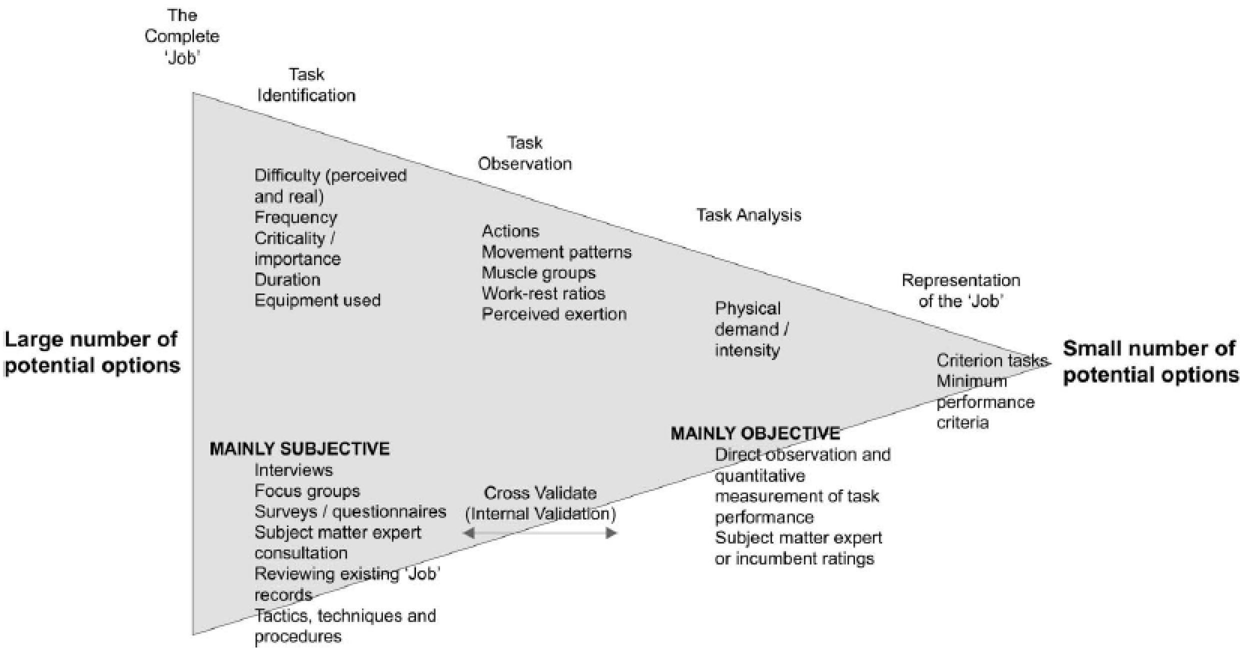


Figure 3-9: Down Selection from all the Physical Job Tasks to Identification of the Critical or Essential Job Tasks [6].

To perform the identification of all potential tasks on a global scale and down select to the *Essential/Critical* ones, there is an analysis of policy, tactical doctrine and training doctrine. This process requires information about task frequency, criticality (consequence), duration, and if there is the need for specific equipment. To collect these data various subjective, qualitative methods are often used such as surveys, interviews, focus groups, questionnaires, SME Consultations, and the reviewing of existing job records [6].

3.3.3 Subjective and Objective Methodology

Surveys and questionnaires facilitate the job task list down-selection process, as they are simple, quantifiable, and can be administered to large samples of incumbents. Whilst it is advisable to conduct large scale surveys across all task domains, it is possible to collect data on smaller samples as long they are stratified to reflect the wider incumbent population [18].

The methods applied to survey the opinions of incumbents might include face-to-face or self-administered approaches that use paper and pencil, online computer platforms or electronic voting systems. The researcher's survey method of choice may be influenced by the access and availability to incumbent personnel, and therefore have a resultant effect on the data quality of responses (e.g., cognitive burden, survey/item response rate, order effects, bias, and willingness to disclose) [19].

Within the PES literature the data quality of subjective job task analysis questionnaires has been explored by embedding known low intensity calibration / control tasks into both urban firefighter (i.e., 'rolling out 38 mm hose', 'using 4.6 m ladder'), and police job task analysis surveys (i.e., 'firearm sight setting', 'strip, clean, assemble firearm') [20], [21]. Analysis of these data could infer response bias and make inferences on the respondents' diligence, and comprehension / understanding of the survey.

When performing a survey or interview in the context of job task analysis it is recommended to ask participants about the physical demands of their tasks and not use general terms such as "difficulty" as Larsen and Aisett [18] indicate that 'difficulty' is an indistinct concept to comprehend and may be open to misinterpretation, and mental demand or skill level may be considered if physical is not specified.

The extent to which we can expect reliable responses from surveys or interviews in unknown as an in depth analysis of the reliability of physical demand as a task domain is yet to be conducted [18]. The stringency / leniency of survey task down-selection criteria does vary [20], [22], and the measure of central tendency applied to generate descriptive data (i.e., mean, median, mode) is debated. If reporting mean, it is recommended that confidence measures be included [18]. It has been suggested that it may be prudent to adopt a lenient / inclusive survey down selection criteria to ensure that potentially relevant tasks are not prematurely excluded by a subjective assessment, at this, an early stage in the PES process [20]. For military PES research in organisations where the "Soldier First" principle has been adopted, a factor analysis, or grouping shared key tasks to identify the underlying factor 'common' to successful job performance across a number of occupations can be employed [18]. It is not recommended to use a "cumulative score" method, by combining measures of the task such as physical difficulty and frequency, as doing so could result in the loss of highly critical tasks which occur very infrequently, and criticality is the most important identifier [5], [18].

The infrequent nature of many military tasks, the potential dangers associated with the observation of some of these tasks, and the inability to quantify certain task domains objectively (e.g., task importance ratings) necessitates the use of subjective job analyses during the Task Identification phase [19]. Unfortunately, this key building block in the PES process is often not fully reported by researchers [23].

This Task Identification phase can often result in the generation of a substantial compendium 20 to over 50 tasks [20], [22], [24], meaning a hierarchical analysis of the *Essential/Critical* nature and physical demands of tasks may assist the down selection of tasks [6]. These final tasks that can be prioritized in subsequent PES steps that involve time, logistical, and cost intensive objective measurements [22].

3.4 SCENARIO CONSTRUCTION

Although job shadowing, or on the job observations are encouraged, Larsen and Aisbett [18] point out that capturing footage of workers performing physically demanding occupations may not allow for the identification of irregular or infrequent yet important tasks. Secondly, objective field data collection may be unsafe for researchers investigating certain tasks (e.g., emergency services and military trades). Third and finally, in less physically demanding occupations, workers may often need to describe more discrete, unobservable tasks, such as balance or fine motor skills, which cannot easily be captured via video [18], e.g., medical procedures or working in confined spaces. For this reason, scenarios and work simulations are also developed with the SME to allow for task reproduction in a controlled field setting. On the job or during training observations are preferred, however in the military setting these critical/essential tasks are often of an emergency or combat nature.

During this Task Identification phase, SME are often involved to help understand the context of each task and to further validate their *Essential/Critical* nature. This process allows the researchers to later reproduce the task in its operational and environmental context, including the state of urgency, threat level, and access to equipment. The SME will aid to determine if each task is actually *Essential/Critical* by discussing the consequences of failure, for example if an incumbent does not achieve a minimum time/load. This process should be facilitated by a trained expert in Focus Group facilitation as the concepts of *Essential/Critical*, and minimum acceptable performance standards are often difficult to comprehend [25].

A subject matter expert has been defined throughout the literature, however SME as related to PES are discussed and defined in the following illustration “Who is a subject matter expert?” (see Figure 3-10) as described by Blacklock et al. [25] and Lee-Bates et al. [14].

3.4.1 Ergonomic Analysis

As military tasks are not often performed on operation as a discrete task, but instead as a part of a whole scenario, the scenarios identified with SME will allow for the physical (forces and loads) and physiological (metabolic demands) measurements to be obtained on a representative sample population [6], [26], [27]. Within these scenarios the participants are both used to measure the demands of the tasks at their operational pace but can also be investigated on their “abilities” to perform these tasks, which helps to determine the legitimacy of a theoretical standard. To be clear, it is not recommended that one uses the performance of the sample to establish the standard alone as this constitutes normative referencing. In some nations this has already been demonstrated as legally indefensible [28].

The scenario generated by SME and reproduced with the research team also allows for the assessment of the influence of external factors (personal protective equipment, environmental exposure, duration and fatigue) on physiological demands. A scenario might last up to an entire day and should reflect work/rest cycles practiced in the field. From these data one can determine the components of each essential task which are the most physically demanding, based on the acquired measurements, thereby concentrating the task; eliminating the redundancies or lower demands; and combining tasks with similar demands. This is the concept of **Combination, Concentration, and Elimination** and is discussed in Section 3.4.3.

When reproducing these tasks and their scenarios a representative population of research participants should include a heterogeneous group of incumbents (variety of fitness levels, age, anthropometrics and sex), and these participants should perform repeat trials to ensure measurement accuracy. For short duration simple tasks such as a box lift, three trials is likely enough to establish familiarity, however, for more complex tasks up to 5 may be needed for performance stability [29], [30], [31], [32].

WHO IS A SUBJECT MATTER EXPERT?

When identifying SME a list of selection criteria have been proposed by Blacklock et al (2015), outlined below, whereby an SME is identified by their ability to meet at least 2 of the 9 criteria. Ideally, there are at least three levels of SME included in the research process: Senior Leaders (commanders/senior non-commissioned members), Mid-level managers (supervisors) and Junior Soldiers (performers) (Reilly et al, 2015).

Criteria for Identifying Subject Matter Experts

1.	Experience performing the task during military exercise or training
2.	Experience performing the task during military deployment domestically
3.	Experience performing the task during military deployment internationally
4.	Experience performing the task during an emergency situation
5.	Experience in a position of leadership where you have directed subordinates to perform the task and have observed the task being performed
6.	Have witnessed the task being performed in an acceptable manner
7.	Have witnessed the task being performed unsuccessfully and can attest to the reasons for, and the consequences of, this failure
8.	Experience witnessing and/or performing the task using several techniques and can comment on the advantages and disadvantages of these techniques
9.	Experience delivering formal training on the task

Lee-Bates et al (2016) investigated if perceptions of physically demanding job tasks are biased by employee demographics and employment profile characteristics including: age, sex, experience, length of tenure, rank and if they completed or supervised a task. These variables were examined in relation to their effects on subjectively defined job task characteristics including ratings of task frequency, duration, distance, physical effort and importance. This research reinforced that SME must include participants who were actively involved in both task participation and supervision for the purposes of rating these tasks and suggests that sex differences may have a small, yet significant effect in relation to perceptions of job performance and ratings of job importance and may be moderated by other participant characteristics [(Landy and Vasey 1991; Van Iddekinge et al. 2005) in Lee-Bates et al (2016) Larsen and Aisbett (2012)]. It is very possible that new recruits perform different duties, or the same duties with differing frequency, to experienced incumbents and that perception of task frequency will vary according to incumbent experience (Larsen and Aisbett, 2012). The criterion for membership at each level of SME input should include job description, current assignment, rank, time served in each critical environment, and how recently the person has served in the position (Reilly et al, 2015). More detail on the requirements of military SME is provided in Reilly et al (2015).

Figure 3-10: Who is a Subject Matter Expert?

Often it becomes difficult to identify a minimum performance standard, in these circumstances videoed performance of the scenario can be used to facilitate SME discussion to distinguish a minimally competent performer from one who is not [16]. Because all SME may not agree, best practice requires a systematic approach that solicits the perspectives of a variety of SME – referred to as a standard setting study.

The ultimate goal of standard setting is to make the resulting minimum cut score as objective and reliable as possible. Thus, documenting the process by which the minimum cut score is established is also critical [33]. Expert panels are asked to judge at what level on the outcome measure a person has failed to meet the minimum requirements of the job. Then they are asked to identify the consequences of false positives and false negatives and to determine what levels of false positives and false negatives are acceptable [25],[33]. This requires that the SME establish a common understanding for what constitutes being minimally qualified [33] which was determined with the definition of Essential/Critical.

For this reason, the re-enactment of the scenario for SME is a good exercise to fine tune the minimum performance requirements, as the SME have the opportunity to complete the tasks, witness others completing the task (live or on video) and may question their original decisions. Or, it may help them to become more specific about determining the standard. An example of this exercise is the CAF Escape to Cover task (Case Study, see Figure 3-11), or the US Army Field Artillery task (Case Study, see Figure 3-12), a tool used to facilitate SME can be as simple as the following table describing levels of certainty for which SME are required to assess the pace of a task for operational effectiveness (pace of carrying sandbags).

CASE STUDY

Canadian Armed Forces: Determination of a minimum time standard to complete an Escape to Cover Scenario

The Escape to Cover task was one of six final essential tasks for the Canadian Armed Forces. SME in the focus group discussed this task extensively and developed a scenario, however to establish a minimum time standard to complete the scenario reenactment in the field was necessary. This task was completed by SME repeatedly at three different paces, too slow (at risk of injury due to live fire), too fast (at risk due to a loss of situational awareness), ideal (as close as possible to the pace they adopted when performing the task in theatre). Their performance was videoed, but their identities concealed, then SME viewed these trials, while voting independently on which pace was the Minimal Operational Performance Standard, similar to `bookmarking` (Lewis et al, 1999). The standard (time to complete) was established once 80% of the group felt that there were no risks associated with the pace.

Figure 3-11: Case Study: Canadian Armed Forces: Determination of a Minimum Time Standard to Complete an Escape to Cover Scenario.

Figure 3-12 provides an example of an SME scale to rate the task of building a bunker.

The SME can be involved in the task observation and can advise on the Method of Best Practice (MOBP) of each task, for example the safe way to lift ammunition. In some instances the MOBP cannot be established by the SME because of variations in practices within an organization [16], or there are multiple safe ways to achieve the task which allow for compensation depending on strengths and weakness. For example, a squat lift may be taught to be the best practice for safety reasons, however if one is physically strong enough to perform a dead lift without injury, this method requires less physiological demand [33]. It has been suggested that this MOBP should be established to determine the most efficient way to undertake a task [34], however one must acknowledge that with a heterogeneous population, the most efficient method will likely be different depending on the individual's physical (such as anthropometry) and physiological characteristics, as well as familiarity with the task. In some instances, not having a MOBP can provide a means of accommodation or facilitate success on the task for low performers. The CAF include a sandbag lifting task in their PES and allow for shorter individuals to assist the lift by supporting the bag with their knee, which is a method of accommodation and facilitates the success of smaller members. The technique associated with task performance may be defined by the scenario, a lift task may have more flexibility when performed solo than one which requires a bag to be passed to another military member. Sometimes legislation requires proper technique be exercised as a duty of care to prevent injury, however there are often exceptions where these are exceeded for those working in Emergency Services [35].

The Case Study from the US Army (Figure 3-12) provides an example of a scale developed for SME to rate the task of building a bunker. During the development of task standards for the recent US Army PES effort, a senior commander required the final standards of performance on each task be verified to ensure that 90% of a randomly selected group of incumbents could perform to the required standard. If 90% of the incumbents were not successful, the standard was re-examined.

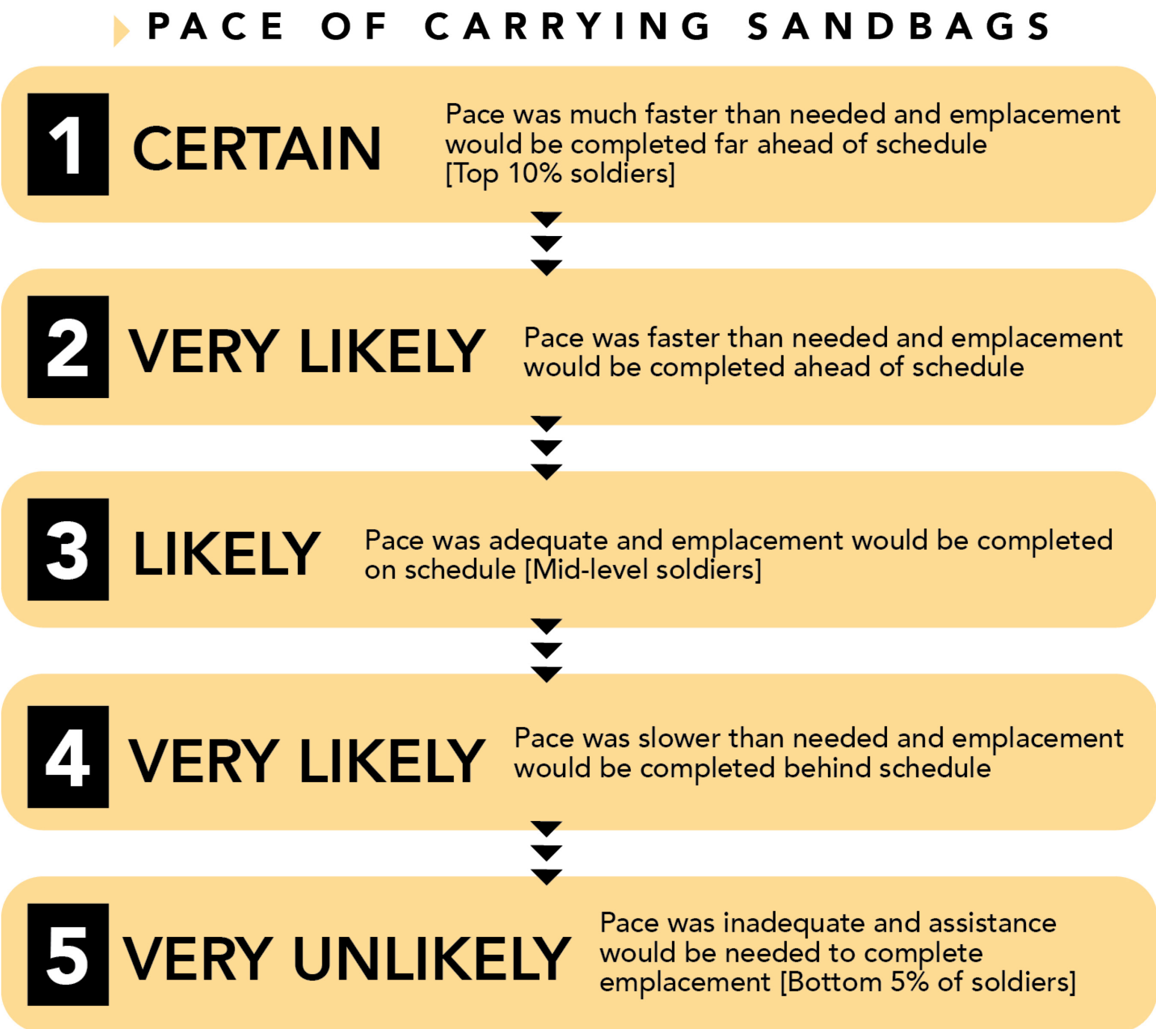


Figure 3-12: Example of an SME Scale to Rate the Task of Building a Bunker.

Other examples of military PES are described in the following three case studies (Figure 3-13; Figure 3-14; Figure 3-15). These three examples provide a variety of methodologies for establishing the minimum performance standards, some of which would not be defensible in Canada, for example, but have proven to be implementable in their country of origin. This is a testament to the appreciation of local legislation, and perhaps a reflection that these methods are yet to be challenged in the court of law.

CASE STUDY

US Army Field Artillery Ammunition Supply Vehicle reloading task

This task requires soldiers wearing 16-23 kg of task specific gear to lift a 48 kg M795 155 mm round from the vehicle tailgate, carry it 3 m and place it into a honeycomb rack with openings from floor to shoulder height For this task, there was no specific doctrine to guide a performance standard. The need for a minimum standard was justified, as the threat of going too slowly leaves soldiers defenseless. The SME indicated the task required two people from a three person crew consisting of a team chief, an ammunition handler and a driver. The chief and the handler were required to load 90 rounds in 30 min. A field training exercise was conducted using active duty Cannon Crewmember soldiers to determine if the standard was reasonable and less than 90% of them were capable of performing the task at that rate. Only 35% of the soldiers tested were able to meet this standard. These trials were videoed and presented back to SME panel, and with this information the SME re-evaluated and agreed that all three crewmembers could participate in the loading task as long as one person stood watch at all times. In addition, the time was increased to 45 min to accomplish the 90 rounds. Subsequent field training exercises revealed that 88% of trained soldiers were able to complete the task to standard.

Figure 3-13: Case Study: US Army Field Artillery Ammunition Supply Vehicle Reloading Task.

The WASL (Working Above Shoulder Level) test of the Netherlands Armed Forces (Figure 3-14) demonstrates the use of various data sources to collect the occupational demands of working above shoulder height through observations, doctrine, and interviews as outlined in the PES development process (Figure 3-1). Both medical and military SME were involved in making the assessment of a minimum performance requirement using scientific methodology and a heterogeneous participant and evaluator sample.

The PES described by the Norwegian Armed Forces is counterintuitive to the guidelines provided in this chapter, since their PES are sex and age adjusted for most positions, and based on norm data (see Figure 3-15). To date these norm-referenced PES have not been challenged in the court of law. This may be due to the fact that the civilian Norwegian Equality and Anti-discrimination Act states that discrimination includes both direct and indirect differential treatment. Indirect differential treatment means any apparently neutral provision, condition, practice, act or omission that results in persons being put in a worse position than others on the basis of sex, age, etc. Thus, age and sex adjusted PES are not necessarily considered discriminative. Moreover, in the Norwegian

Armed Forces differentiated PES allows for selecting the most physically fit men and women. Neutral PES will lead to:

1) Reduced requirements for men, or

2) Very few women will pass the requirements.

CASE LAW: NETHERLANDS ARMED FORCES (NAF)

Working above shoulder level

Introduction

In the recent revision of the NAF pre-entry screening, working above shoulder level (WASL) was recognized as one of the 25 essential military job components to be translated into a medical capacity standard. A study by the Army Sports Medical Centre and the Dutch Coronel Institute for Occupational Health developed a reliable functional pass-or-fail test for WASL, including assessment criteria that produce 0/1 dichotomous variables, to be used in the NAF pre-entry screening.

Methods

To develop the WASL test, military employees from units of all three operational commands (Army, Navy, Air Force) who performed tasks involving WASL on a regular base were systematically assessed (live observations, videos, expert interviews, internal reports) and analyzed (multi-criteria analysis). The following descriptive variables were used to quantify work performance: duration, frequency, posture, precision demands, and force magnitude. These data were used to make an accurate translation of WASL-related military activities into a functional test.

Following the test development phase, intra- and inter-observer reliability were determined by three physiotherapists (two males, one female) working at the NAF Pre-entry Screening Center score videos (n=37) of a subset of NAF military personnel (n=28; 24 men, 4 women). This is representative of the pre-entry screening population (based on age, sex, and education) performing the WASL functional test. The same videos were scored a second time in random order five days later.

Test design

Test Set-up and materials

The test takes place in an indoor sports facility. A wooden box (width 58 cm, height 20 cm, depth 26 cm, with left and right side of this box holds two round openings left and right, diameter 6 cm) is hanging with four metal hooks at two rungs of a wall rack. The openings are 24 cm deep, 29.5 cm apart from each other, and 8.5 cm above the bottom of the box. A round, oblong weight of 6.1 kg just fits in two tubes. Between the two openings, 2 bolts can be placed, 10.5 cm from the openings and 10 cm from the bottom of the box. The bolts have a length of 6.5 cm and a diameter of 10 mm. A butterfly nut can be screwed onto each bolt.

The height of the box is adjusted to the height of the subject, in that the height of the bottom of the box equals the height of each subject. Hereby, the working level (and, therefore, the degree of difficulty) is equal for each subject.

Figure 3-14: Case Study: Netherlands Armed Forces (NAF) – Working Above Shoulder Level.

CASE LAW: NETHERLANDS ARMED FORCES (NAF)

Working above shoulder level

Test protocol

Before the test, the test leader (physical therapist) gives instructions on how to perform the WASL functional test. During the test, no verbal instructions or support is given.

At the start of the test, the applicant needs to put a 6.1 kg weighing, round, oblong weight in the left opening in the box; the butterfly nuts need to be completely screwed onto the bolt. The subject stands in front of the box. The horizontal distance between the subject and the box is not fixed, but depends on personal preferences of the subject (in accordance with field observations). Time starts with an audible beep when the subject touches the weight. Every 4 seconds another beep will be heard (14 beeps in total). At every beep, the subject moves the weight (using both hands) from the first opening to the other, which means moving the weight forth and back 15 times in total. Beside repetitive movement, this part of the test also represents gross precision demands and heavy manoeuvring.

After the subject has moved the weight 15 times, a double beep sounds (one minute after the start of the test), indicating that the subject has to start unscrewing and screwing the butterfly nuts for one minute. This activity represents the fine precision demands and static posture. After this minute, another beep sounds and the test is finished. During the performance of the test, the subject is not allowed to lower his/her hands below the bottom of the box or to pause for longer than 2 seconds.

To pass the test, the subject has to meet the following criteria:

I. During the test, the hands of the subject must be at or above the bottom of the box.

II. In the first part of the test, the subject has to move the weight each time the beep sounds.

The subject is allowed to miss one beep at the maximum, but has to catch up this beep within the first minute of the test, so the subject moves the weight 15 times in total.

I. The subject should not drop the weight.

II. The subject should not pause for longer than 2 seconds during (un)screwing the butterfly nuts, even if his/her both hands are above shoulder level.

The test leader assesses whether the subject meets all these criteria.

Figure 3-14 (Continued) Case Study: Netherlands Armed Forces (NAF) – Working Above Shoulder Level.

This is due to the natural physiological differences between sexes. Excluding women from many military positions is not considered acceptable, neither is reducing male soldiers' PES. In addition, it would be a massive undertaking to analyse task and demands for all positions (more than 100 positions just for conscripts). When the resources do not allow for solid task and demand analysis, the Norwegian Armed Forces consider it potentially unethical and discriminative to set arbitrary neutral standards. Thus, the Norwegian Armed Forces primarily uses a norm-based approach with sex and age adjusted requirements. This is also in part related to the fact that Norway do not have Universality of Service or the Soldier First principle.

CASE STUDY

Norwegian Defense Force

A norm-based system was used to develop the PES-scales for all the predictive tests. The 1-9 minimum requirement (MR) scale is based on test data collected among male and female conscript recruits between year 2010 and 2014. For men, the performance corresponding to the 5th percentile equals MR 1 (low physical demands), while the 80th percentile equals MR 9 (high physical demands). For female soldiers, MR 1 is also at their 5th percentile score, while MR 2-8 are more strict than for men. A similar approach is used to adjust the PES for age. The absolute test score requirement for MR 9 is similar for all, irrespective of age and gender (i.e. typically at ~99th percentile for women).

Each branch (Army, Air Force, etc.) sets their own minimum standard levels for their different trades, within the abovementioned 1-9 scale. Accordingly, each branch carry out their own physical demand analysis, and decide their own minimum requirement. The type and quality of the physical demand analysis vary between branches.

Figure 3-15: Case Study: Norwegian Defence Force.

The strength tests previously used in the Norwegian Armed Forces were push-ups, sit-ups and pull-ups. Push-ups and sit-ups have in previous studies been shown to be poor predictors of occupational physical performance [36]. Although the pull-ups test was retained, push-ups and sit-ups were removed in the new set of physical fitness tests, and replaced by broad jump and standing medicine ball chest throw [37]. Previous studies have found broad jump to be a good predictor of lower body explosive power [38], [39]. It is also a field expedient test requiring limited equipment and has a low level of technical difficulty reducing the test-retest variability. Standing medicine ball chest throw was recommended by NATO as a field expedient test of maximal and explosive upper body strength [40], and medicine ball chest throw performance have been shown to be highly correlated with upper body strength [41]. Both of these tests were validated against performance on a casualty drag and 1 repetition maximum (1RM) bench press, and it was shown that broad jump correlated well with casualty drag performance (r = -0.73), and that medicine ball chest throw distance correlated well with 1RM bench press (r = 0.84) [42].

The RAF Regiment Battlefield fitness test provides an example of how tasks should be simulated within the constructs of their operational scenario and not always as discrete tasks. Although the method of establishing the minimum standard (time to complete) was determined by norm-referencing at the 95th percentile, meaning that only 5% of incumbents would fail, the researchers did provide evidence that those who did not pass, validated their exclusion, and some elements were determined to be successful if completed continuously without a standard.

When designing PES, it is important to avoid standard "creep". The "more fitness is better" theory has no bearing on the minimum standard or cut-off necessary to perform the task/job successfully. For this reason, ranking employees on their performance on job simulations and predictive fitness tests is not a valid method of determining which applicants are best for the job.

3.4.2 Physical and Physiological Measures

While performing the scenarios and tasks in their operational setting, or simulation of, it is advised to use this opportunity to measure the physical (biomechanical constructs) and physiological (metabolic constructs). This measurement phase typically requires considerable scientific equipment such as portable metabolic measurement devices, accelerometery, or load cells. Before deciding what system or combination of systems to use some considerations include:

1) Interaction of the device with personal protective equipment (e.g., heart rate monitors can be painful when pressure is exerted by torso-borne ballistic protection).
2) Duration of the task, and sampling frequency (some military exercise last weeks).
3) If the exercise is whole body or segmental analysis is required.
4) Energy system.
5) What is the mechanism for task failure (does one fail a heavy lifting task due to strength or aerobic abilities)?
6) The potential for the measurement device to alter the integrity of the task (attaching a load cell to a casualty as opposed to gripping the casualty with one's hands).

These data will facilitate the determination of the components of fitness required for successful task performance.

3.4.2.1 Components of Fitness

The definition of military physical fitness is very broad, and it is generally accepted that physical fitness is the result of a number of specific types of behaviors, attributes and capabilities that are termed "components" of physical fitness [43]. The British Army defines ten distinct components of military physical fitness (Figure 3-16), which underpin performance involving any physical activity.

A comprehensive literature review that summarized factor analytic studies in combination with physiological studies to determine the major components of military physical fitness for the United States (US) Army identified similar components of physical fitness [43]. The major components identified were cardiorespiratory (aerobic) endurance, muscular strength, muscular endurance, flexibility, coordination and balance.

One can break down the tasks within the scenario theoretically hypothesizing which components of physical fitness are necessary for success on the task. This is where attention must be paid to the definition of task success, is it more important to rescue a casualty fast or to be able to rescue a casualty with a higher than average mass, or rescue a casualty single handed as opposed to a part of a team? One measure of success may depend more on muscular endurance and aerobic fitness whereas the others may depend more on strength.

For example, the CAF research team measured the metabolic demands of a scenario which involves a post-earthquake humanitarian mission. CAF members were required to remove rubble and search for human casualties, then perform the required casualty evacuation. This task was simulated for four hours and was not performed at a high pace or high intensity therefore respecting the work/rest cycle policy. Once these types of scenarios and tasks have been performed, and the physiological demands captured, the list of tasks is reduced to the final core tasks and competencies, through down selection (Figure 3-17) or **concentration, combination, and elimination**.

NATO HFM RTG 269 Combat Integration: Implications for Physical Employment Standards

Fitness	Fitness Component	Definition	Example Activities
Endurance	Aerobic Capacity	Ability to sustain sub-maximal low-to-moderate/high intensity activity for a sustained period of time (minutes to hours), typically involving dynamic whole-body activities	Sustained patrolling carrying load (e.g. ≥ 30 kg) or digging a fire trench
	Anaerobic Capacity	Ability to sustain intermittent or continuous near maximal or maximal efforts for a short period of time (seconds to minutes), typically involving dynamic whole-body activities	Fire and movement task or a break contact task
Strength	Muscular Strength	Ability of a muscle group to exert maximal force in a single voluntary contraction (< 5 seconds)	Lifting objects, e.g. a casualty, equipment onto a vehicle. Standing up from kneeling while carrying a heavy load
	Muscular Endurance	Ability of a muscle group to repeatedly generate an intermittent or continuous moderate-to-high absolute force for a more prolonged period of time (seconds to minutes)	Repetitively lifting and carrying stores or a stretcher casualty evacuation
	Muscular Power	Ability to exert maximal external force in the shortest possible time (typically less than 1 second)	Breaking down a compound/ building door or jumping over a ditch or low wall
Mobility	Flexibility	The ability to voluntarily stretch, flex or lengthen parts of the body as far as possible i.e. the range of motion around a joint	Lifting a leg over a fence or bending down to pass under a low obstacle
	Balance	Maintenance of equilibrium while stationary or moving	Maintenance of a stable firing position
	Speed	Ability to perform movements in a short period of time	Rapid movement between fire positions
	Agility	Ability to change the position of the entire body in space with speed and accuracy	Hurdling a fence or rapidly changing running direction (e.g. fire and movement task)
	Coordination	Ability to synchronise the senses (e.g. sight/hearing) with body parts to move smoothly and accurately	Bringing weapons systems to bear and accurately engaging with the enemy

(Blacker S. et al, 2018)

Figure 3-16: Components of Fitness.

3.4.3 Concentration/Combination/Elimination

This method is similar to the down-selection process described in Reilly et al. [6] (Figure 3-17). The process requires those with knowledge and understanding of the components of fitness as well as energy systems to evaluate the physiological data obtained on each task and determine which components of each task are the most demanding and supersede those which are less consequential. The idea being that if one has the physical capability to perform the resulting task (in this case lifting 60 sandbags) then one will have the fitness required to perform the other roles in the task scenario as well as tasks with similar or overlapping demands. **Concentration, combination, and elimination (**Figure 3-17**)** demonstrates how the demands of both building a sentry post and sandbagging for a flood can be reduced and concentrated to the highest demanding component of lifting the sandbags. When employing the work/rest cycles as per military doctrine, which were re-enacted during the scenario, other roles such as passing sandbags are of lesser or negligible demand. Often researchers are asked about the fitness required to complete a task over a long duration. These scenarios provide the opportunity to evaluate this concept and determine if the demands increase over time require a higher level of fitness, or if the work/rest cycle facilitates sufficient recovery. In the case of this sandbagging task, incumbents of varying fitness levels were used as research participants and all of them achieved resting or low levels of energy expenditure on the rest phase of the work: rest cycles and did not demonstrate any cumulative fatigue after performing this task for a long duration.

3.5 TEST DEVELOPMENT

At this point in the research process the PMT and the research team can decide if they will proceed to develop a task simulation style test, a predictive fitness test, or a hybrid of both. This will be determined by the balanced need for both Fidelity and Feasibility (Figure 3-18) we will discuss in this section. Whatever the decision, at this phase in the research process the PMT should meet, review the final list of tasks (the most physically demanding) which will be reflected in the final PES, and weigh in on the concept of Fidelity and Feasibility to ensure that the final outcome can be implemented and delivered given any constraints on the organisation.

This phase of the research process can be the most demanding, as it includes large scale trials, meaning many measures on a large sample. The Essential/Critical tasks and their defensible minimum standard have already been established and the most physically and physiologically demanding tasks have been determined through direct measurement and Combination/Concentration/Elimination resulting in the understanding of the Components of Fitness necessary for safe and successful performance on the tasks. The researcher and PMT can investigate the relevance of a predictive style test, a task simulation or hybrid (both) throughout these trials involving participants performing both the Essential/Critical task simulations and the potential simulations/tests to maximum volitional fatigue.

Before deciding on predictive, simulation or hybrid there are many considerations to make with the PMT to determine what style test is best for the resulting PES based on predictive accuracy, validity (content, logical, criterion, and construct) [16], simplicity and manageability of the potential PES models [6].

3.5.1 Test Design: Balancing Fidelity and Feasibility

An important consideration when identifying potential screening tests is the balance between **Fidelity and Feasibility** [44]. Fidelity to the job refers to the similarity between the test and job tasks; whereas feasibility highlights the functionality, operability, and reproducibility of results across test location and facilitator. Naturally, these two concepts anchor opposite ends along the continuum as tests, with high fidelity typically lacking feasibility and vice versa. While high fidelity job simulations are reflective of physical demands with obvious overlap to job tasks, they are costly to administer, difficult to replicate, and not generalizable across multiple jobs [44]. The simulative nature of high-fidelity tests offers several benefits such as *face validity* (the perception of fairness by participants) and legal defensibility (with the appropriate supporting job analyses). However, large organisations (i.e., military) face unique challenges in cost and replicability of administering high fidelity tests across diverse jobs and wide-spanning locations. Instead, administering a series of inexpensive, easily replicable, predictive fitness tests that could be generalized across jobs may be a more feasible approach. This style of testing tends to lack specificity, face validity, however it remains to be seen if it is more susceptible to bias. PES can be categorized anywhere along the continuum in that the most successful assessments demonstrate components reflecting both **Fidelity and Feasibility** Figure 3-18.

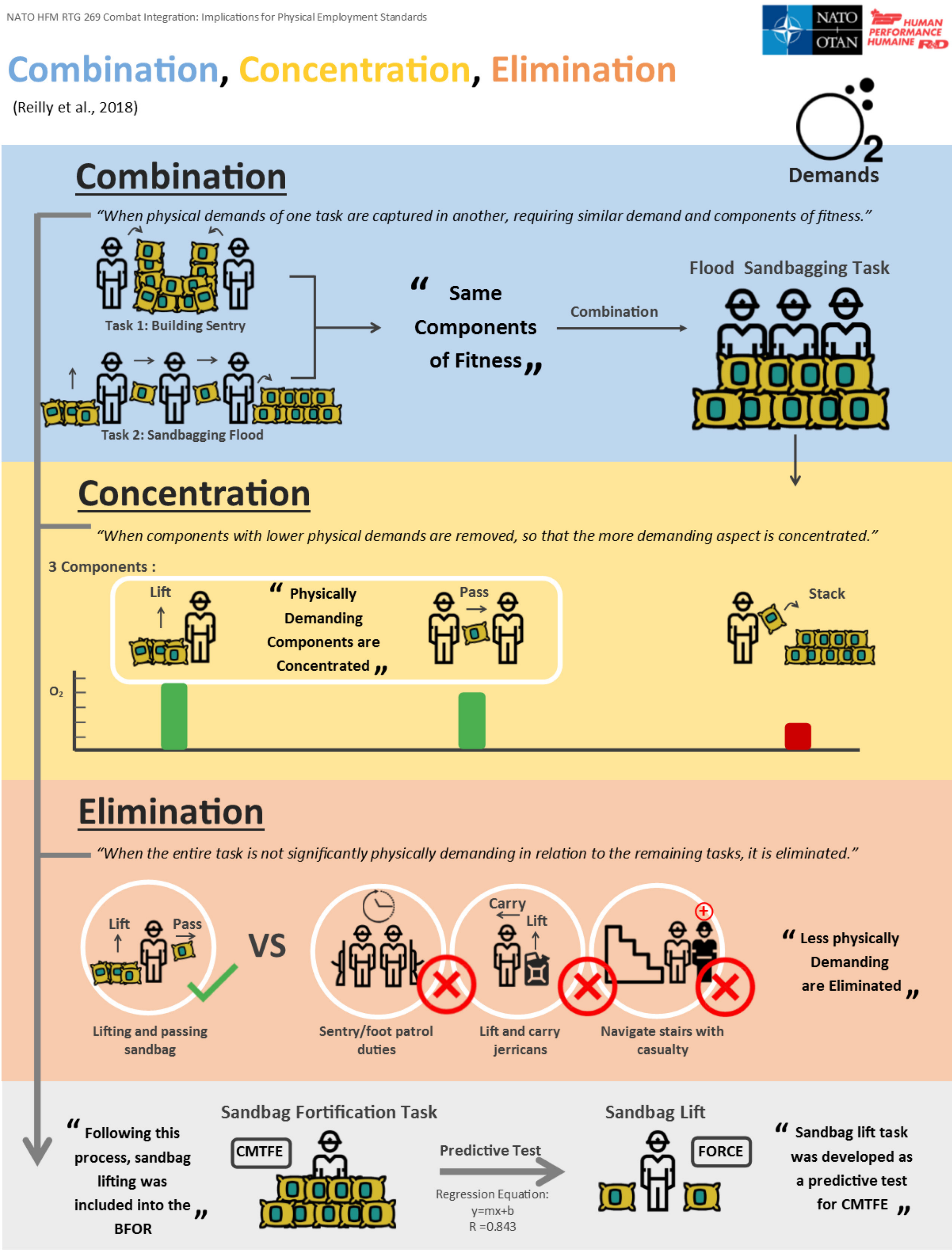


Figure 3-17: Concentration, Combination, and Elimination.

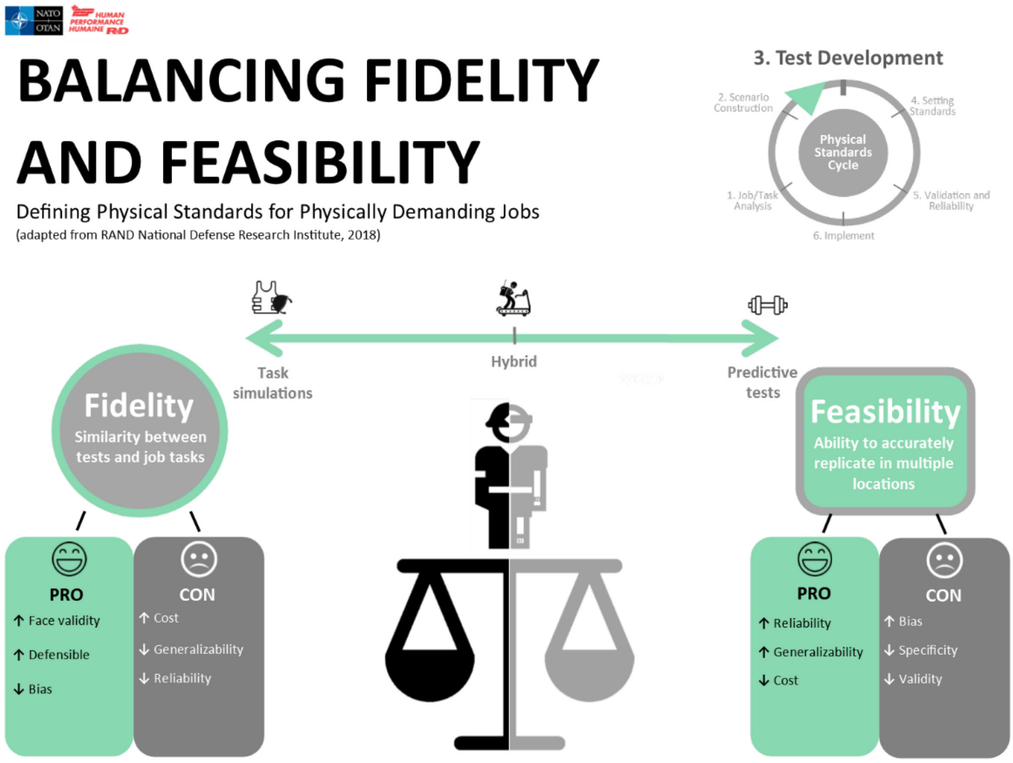


Figure 3-18: The Balance Between Fidelity and Feasibility.

As an example of feasible testing with adequate fidelity, the CAF FORCE Evaluation consists of 4 predictive occupational fitness tests, which were developed based on the physical requirements of the task simulations (the BFOR of CAF). The CAF member has three attempts to pass the predictive tests (FORCE) and one attempt at the task simulations (CMTFE) before being determined non-operational and dismissed. The FORCE test is a predictive test that takes 30 minutes and therefore is highly feasible while the CMTFE is a task simulation test that takes all day but has high fidelity.

A task simulation test directly replicates an essential/critical job task, in contrast to a generic predictive test which assesses a physical or physiological capacity or construct, or **Component of Fitness** [5]. Aside from having high content, criterion and face validity, a key attribute of task simulation tests is that their predictive validity is immediately apparent given the logical relationship between the task simulation test and the corresponding criterion (essential/critical) task [6]. Generic predictive tests, although feasible and easy to administer, lack face validity, and a single physical construct (i.e., aerobic capacity) may account for only a small proportion of the overall task performance [5], [6]. For this reason, a battery of predictive tests may be needed to test all the important physiological requirements of the task, as humans have the ability to compensate energy systems or muscles recruited to a certain degree, for example, success on a lifting task can be performed in multiple ways to achieve the same goal, not simply upper or lower body strength.

Predictive tests do not require any job skills and may be safer and more appropriate for screening potential recruits into a job for which they have the physiological capacity, but lack experience. Historically, military organisations have relied upon fitness tests (e.g., timed runs, push-ups, sit-ups) as they easily apply to mass testing, require little to no equipment and are faster for participants to undertake, therefore highly feasible. However, these tests have been met with much criticism as they are typically not strong predictors of occupational task performance and lack face-validity [45]. However a system of predictive tests, when validated against the criterion tasks, have the power to discriminate and select successful workers, while not increasing sex bias [46], [47]. In addition, performance on these predictive tests is usually compared or qualified with normative data, which is typically sex and age normalized, and this creates confusion when the standard should be sex and age free. The pitfalls of using the historical predictive fitness tests like grip strength, push-ups and bench press as they relate to bias and combat integration will be discussed in Chapter 5.

If a military can support a task simulation style test then task reproduction and simulation should require relatively low levels of skill learned in training or on the job otherwise participants should be familiarised to ensure that skill level will not have a major influence on task success [6]. There is always a risk that the simulation does not accurately replicate the task, and research has identified that researchers should avoid simplifying complex tasks when reproducing field posture exposure in laboratories, since omitting extra subtasks may lead to an inaccurate reproduction of field exposure [5], [48]. In summary, often a highly feasible mass testing style predictive test is applied to applicants, but a task simulation test is applied on more experienced incumbents or occupations such as Special Forces where smaller groups are being tested and fidelity needs to be maximised.

3.5.2 Critically Analyse for Bias

Before conducting the large-scale trials, the study must have sufficient statistical power (i.e., a large enough number of test subjects) to obtain a reliable estimate of the relationship between the predictive tests and performance on job simulations. The power calculation should incorporate the best information available within the organisation or from external sources on the expected distribution of performance scores on the tests. Data in Chapter 5 can aid in these calculations. The power calculations may show that some population subgroups (such as women or other groups) should be oversampled [44]. In cases in which groups are oversampled, complex sampling statistics need to be applied in the subsequent analyses to determine if these subgroups can be combined with the mass population to facilitate ‘one test’ for all. To ensure the best possibility for success one must approach the test design knowing the biases inherent in each type of measure (discussed in Chapter 5) so that the potential for a sex-free PES is maximised.

3.5.3 Measure Maximum Performance to Establish Correlations

Once the sample size and characteristics are determined, the test battery is developed for both the task simulations and their predictive tests. A protocol must be designed to ensure that participants can demonstrate a maximum effort on all components to ensure that continuous data is captured, permitting the use of regression analysis. If maximum effort is not achieved, a predictive relationship between task and test cannot be assessed [49]. As many task simulations have no previously established safe protocol to maximum effort (e.g., casualty extrication), participants are often instructed to demonstrate their “safe ceiling” or “safe maximum” by gradually increasing the load or tempo, through repeated trials, whilst trying not to induce unnecessary fatigue.

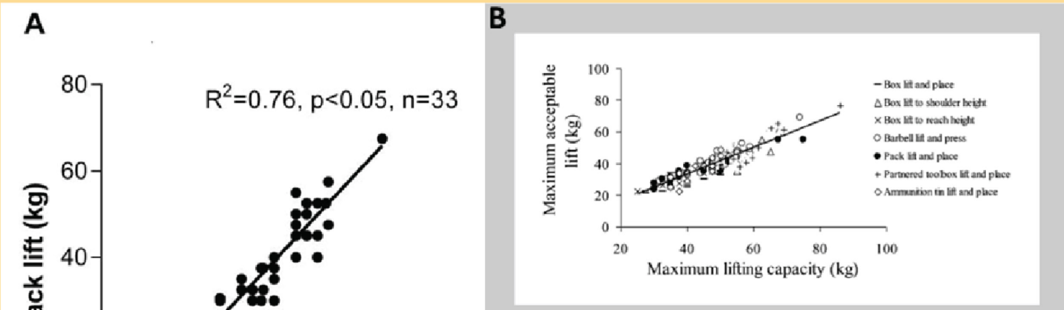
It has been suggested that one should avoid taking people to a 1RM and instead take them to their 5 RMs (USAF) using this as their maximum measure of performance on the task or test (Personal communication, Neal Baumgartner). This is a viable option if the task is to be repeated in the operational scenario, however if the task would only likely be performed once at maximum in the field it may not give an accurate understanding of one’s true operational abilities. Ideally, PES testing should be designed to allow maximum effort to volitional fatigue. If not, the potential applicant could argue that they would have passed the test had they been given the opportunity to exert maximal effort, and submaximal data on the task simulation is difficult to correlate with maximal performance on the predictive tests.

The relationship between what is perceived to be a Maximal Acceptable Weight of a Lift (MAWL) and the Maximum Lifting Capacity (MLC) was presented by the Australian Defence Force scientists and described in the Case Study: Australian Army: Setting a safety margin on a strength-based task (Figure 3-19). This work demonstrates that if participants are to lift to their “safe” maximum, they may have 18% reserve capacity.

CASE STUDY

Australian Army
Setting a safety margin on a strength-based task

The box lift and place PES assessment is a task-related predictive test that assesses muscular strength capacity in Australian Army personnel (Carstairs et al., 2016). The box lift assesses the ability of personnel to lift a field pack into the back of a common military vehicle, which is deemed an essential task for personnel (Carstairs et al., 2016). Based on the relationship between maximal pack lift and maximal box lift performance (r^2=0.76, Figure A), test standards were established. In meeting these standards however, it is unknown if personnel have any reserve lifting capacity above these standards. It is important to not just have the physical capacity to lift a pack (or other objects), but have the capacity to lift the pack safely. To determine this, soldiers' maximum lifting capacity (MLC) and maximum acceptable weight of lift (MAWL) was assessed across seven common military lifting tasks (Savage, Best, Carstairs, & Ham, 2012). The MAWL is a psychophysical measure that represents a tolerable self-selected mass that favours the comfort and safety of the lifter, as opposed to their MLC, reducing the risk of injury. Across the seven criterion lifting tasks it was observed that MAWL was 84% of MLC (Figure B). Therefore 19% was added to the original box lift standard to ensure that personnel are capable of performing the criterion lifting task safely. Beyond a decreased risk of injury, it also increases the likelihood of task completion in a fatigued state (e.g. during operations/deployment).



Pooled maximum acceptable lift and maximum lifting capacity data for 7 strength-based lifting tasks. Data from 7 individual strength-based lifting tasks were combined to develop a generalized relationship between maximum acceptable lift and maximum lifting capacity across a broad range of lifting tasks. From linear regression, if the trend line were to intersect each axis at zero, then a simplified relationship can be expressed in the following form: Maximum acceptable lift = 0.84 × (maximum lifting capacity).

Figure 3-19: Australian Army – Setting a Safety Margin on a Strength-based Task.

3.6 SETTING STANDARDS

At this point in the process we have already established the minimum acceptable standards on the tasks, but if a prediction with a test is involved, such as those assessing **components of fitness**, now a standard on the test must be established. Using the maximum performance data obtained from the large-scale trials, one can determine if acceptable power was achieved, and there are a variety of ways to analyse and compare the relationships between performance on the essential/critical tasks and their predictors, regression and correlation being the most frequently used. For a linear relationship, these analyses use the least squares, best-fit approach to minimize the sum of the squared residuals. The R value is often reported and the R^2 (coefficient of determination), is an estimation of the extent to which variability within the dependent variable may be explained on the basis of variations observed within the independent variables. However, the criticism to this approach is that causality is implied, increase performance on the predictor (run speed) and the task performance will improve (perhaps dig speed). This is erroneous as a R^2 indicates a relationship but not causality. A statistically significant relationship does not imply physiological significance or importance [50].

Typically, Ordinary Least Squared (OLS) regression is used to determine the relationship between two variables. However, in contrast to OLS regression [50], Ordinary Least Product (OLP) regression assumes random error is present in both the dependent and independent variable, producing a common linear equation that gives identical results when predicting y on x and x on y [50], [51]. This is an important distinction from a practical perspective, which Ref. [52] demonstrated for the relationship between a 2.4 km run time and performance on the Multistage Shuttle Run Test (MSRT). Both these field-based running tests are used to measure aerobic fitness and are used interchangeably within the military. If OLS regression had been used to derive the estimated MSRT score for an individual who had just run 2.4 km in 13:00 min:sec (using Equation (3-2)), the result would be 58 shuttles (level 7 and 8 shuttles). If OLS regression was used to derive the estimated 2.4 km run time for an individual who had just achieved 58 shuttles on the MSRT (Equation (3-3)), the result would be 31 seconds quicker at 12:29 min:sec.

This makes the choice of the dependent variable in the ordinary least squares regression analysis of critical importance, and the difference between the two methods will become increasingly divergent as the dependent variable score moves further from the mean value of the chosen dependent variable. Using OLP regression (Equation (3-1)), 58 shuttles would equate to a 2.4 km run time of 12:43 min:sec, irrespective of the choice of the dependent variable.

MSRT performance (total shuttles)

$$= (9.702 \text{ x } 2.4 \text{ km run speed } [\text{km} \cdot \text{h}^{-1}]) - 52.56 \quad (3\text{-}1)$$

MSRT performance (total shuttles)

$$= (8.879 \text{ x } 2.4 \text{ km run speed } [\text{km} \cdot \text{h}^{-1}]) - 40.95 \quad (3\text{-}2)$$

r = 0.92 ($p < 0.01$), SEE 7.7 shuttles

$$\begin{aligned} &2.4 \text{ km run speed } (\text{km} \cdot \text{h}^{-1}) \\ &= (0.094 \text{ x MSRT performance [total shuttles]}) + 6.15 \end{aligned} \quad (3\text{-}3)$$

r = 0.92 ($p< 0.01$), SEE 0.79 $\text{km}\cdot\text{h}^{-1}$

$$\begin{aligned} &\text{MRST performance (total shuttles)} \\ &= (9.913 \text{ x } 2.4 \text{ km run speed } [\text{km} \cdot \text{h}^{-1}]) \\ &+ (5.32 \text{ x Sex}) - 57.34 \end{aligned} \quad (3\text{-}4)$$

r = 0.92 ($p < 0.01$), SEE 7.5 shuttles

For sex, male = 0 and female 1

Recommended statistical methods in PES development have been recently reviewed [16], [53].

It is clear that statistical imprecision and/or misuse can directly influence the development and application of PES. The required level of precision should be considered up front and will depend on things such as the measurement tool, and imprecision in the methodology. Petersen et al. [54] provides a good illustration of the need to define the degree of precision or minimal detectable change, before proceeding with a statistical analysis. This was thoroughly debated for the oxygen demands required for Fire Fighters (Case Study: Fire Fighters and the Prediction of Oxygen Consumption, see Figure 3-20).

CASE STUDY: FIRE FIGHTERS and THE PREDICTION OF OXYGEN CONSUMPTION

For this measure a group of 25 scientists (14 universities, 11 government institutes: 382 years of experience in occupational physiology) met and established the threshold for the acceptable limits of agreement (±0.24 L·min−1). With this minimal detectable difference metric determined, the strength of the predictive tests could be evaluated. Results indicated that the predictive tests were found to overestimate the oxygen consumption of those field tasks to levels that exceeded the measurement precision required by experienced physiologists.

Figure 3-20: Case Study: Fire Fighters and the Prediction of Oxygen Consumption.

This step, determining the degree of precision or minimal detectable change is a recommended but often ignored practice for PES research before beginning statistical analysis and evaluation of the data. A similar approach was undertaken for the CAF and the precision of both the instrument and the assessor (human) was identified, in order to contextualize the minimal detectable change, and determine reliability measures of the PES [32].

The predictors for various models on CAF data were selected by performing a correlation matrix (Table 3-1) on maximum performance data from six job simulations (CMTFE) and 13 potential Predictive Selection Tests (PST).

Once the predictors were identified for any given task simulation, they were entered into a regression, seeking out the highest coefficient of determination (see Table 3-2 and Table 3-3).

For example, for the task *Escape to Cover*, (see Figure 3-21) the most significant predictor was found to be the *20 m Rushes*. The regression model of *20 m Rushes* to *Escape to Cover* (to be completed in 68 seconds or less) indicated that the standard for *20 m Rushes* should fall at 49 seconds.

The final PES for the CAF consisted of four predictive tests, all of which could be correlated to performance on *Escape to Cover*, however the adopted minimum standard for *20 m Rushes* was selected as the most highly correlated to *Escape to Cover* and for which the *20 m Rushes* standard was the most difficult to attain. The final standard time required by *20 m Rushes* was determined by power and specificity calculations performed for all predictive models using *20 m Rushes* between 40 and 60 seconds.

It was determined, based on Table 3-3, a standard of 51 seconds on the *20 m Rushes* would be the most appropriate, for successful performance on *Escape to Cover*. At 51 seconds, power and specificity were both within an acceptable range, and Adverse Impact was eliminated (at 51 seconds, women's pass rate was 83.8% of the overall pass rate).

This same process of determining fidelity, was repeated for all six of the task simulations, using each one's most significant independent variable as a basis for power/specificity calculations.

Table 3-1: Pearson Moment Correlations Between CMTFE and PST.

	EscapeToCover Sqrt	Picking Log10	Digging Log10	PicketsAndWire Inverse	Stretcher Carry	Vehicle Extrication	SandbagFort Inverse
LoadedShuttles Inverse	-.647**	-.536**	-.527**	.824**	.530**	.544**	.671**
Rushes Log10	.671**	.432**	.394**	-.684**	-.437**	-.444**	-.562**
SandbagLift Inverse	-.648**	-.669**	-.616**	.631**	.737**	.668**	.843**
SBPull	-.497**	-.631**	-.580**	.494**	.759**	.749**	.694**
FarmersCarry Log10	-.395**	-.537**	-.532**	.509**	.619**	.611**	.606**
Grip	-.421**	-.601**	-.519**	.418**	.657**	.523**	.611**
MSRDistance Log10	-.529**	-.414**	-.368**	.664**	.420**	.393**	.529**
Pushups Sqrt	-.583**	-.506**	-.497**	.590**	.577**	.553**	.599**
YoyoDistance Sqrt	-.569**	-.379**	-.425**	.778**	.428**	.445**	.542**
WallSit Log10	-.188**	-.112*	-.121**	.356**	.161**	.157**	.222**
BackEndurance Sqrt	.001	.090*	.047	.247**	-.101*	-.042	-.003
FlexedArmHang Sqrt	-.458**	-.341**	-.368**	.593**	.398**	.364**	.462**
Plank Log10	-.313**	-.276**	-.288**	.501**	.246**	.256**	.356**

Table 3-2: Coefficients of Determination for Each of the six Regression Models Using Fitness Predictors and a Hybrid Test.

CMTFE Element	Coeff of determination (r^2)		
	Best fit Model (Rushes, Loaded Shuttles, Sandbag Pull, Sandbag Lift,	Fitness Component model. (Push ups, Mod CF Aer Test, Handgrip)	Hybrid model (Rushes, Sandbag Pull, Sandbag Lift, Mod CF Aer Test)
Escape to Cover	0.702	0.524	0.683
Picket and Wire Carry	0.711	0.646	0.705
Sandbag Fortification	0.780	0.574	0.766
Picking and Digging	0.618	0.458	0.537
Stretcher Carry	0.676	0.546	0.671
Vehicle Extrication	0.618	0.415	0.619

Table 3-3: Sensitivity and Specificity Calculations for the Prediction of a 68-sec Standard on Escape to Cover (E2C) by 20 m Rushes.

Time on Rushes (sec)	Pass Rate (%) Overall	Male	Female	Power (%) True pos/True pos + False neg The likelihood of failing the test if one failed the task The proportion of failures that we are able to detect	Specificity (%) True neg/true neg + false pos The likelihood of passing the test if one passed the task The proportion of passes that we are able to detect.
46	70.6	81.9	46.1	91.7 (i.e., 8.3% of failures on the E2C task go undetected by our test)	79.0 (i.e., 21% of failures on our test will likely pass the E2C task)
47	75.3	85.6	53.3	88.3	83.8
48	78.0	86.9	58.7	85.0	86.2
49	81.4	90.0	62.9	76.7	89.1
50	83.3	91.7	65.3	75.0	91.0
51	86.5	93.1	72.5	63.3	93.2
52	87.7	93.6	74.9	56.7	93.7
53	91.1	95.6	81.4	45.0	95.9
54	92.4	96.7	83.2	38.3	96.5
55	93.5	97.2	85.6	35.0	97.4

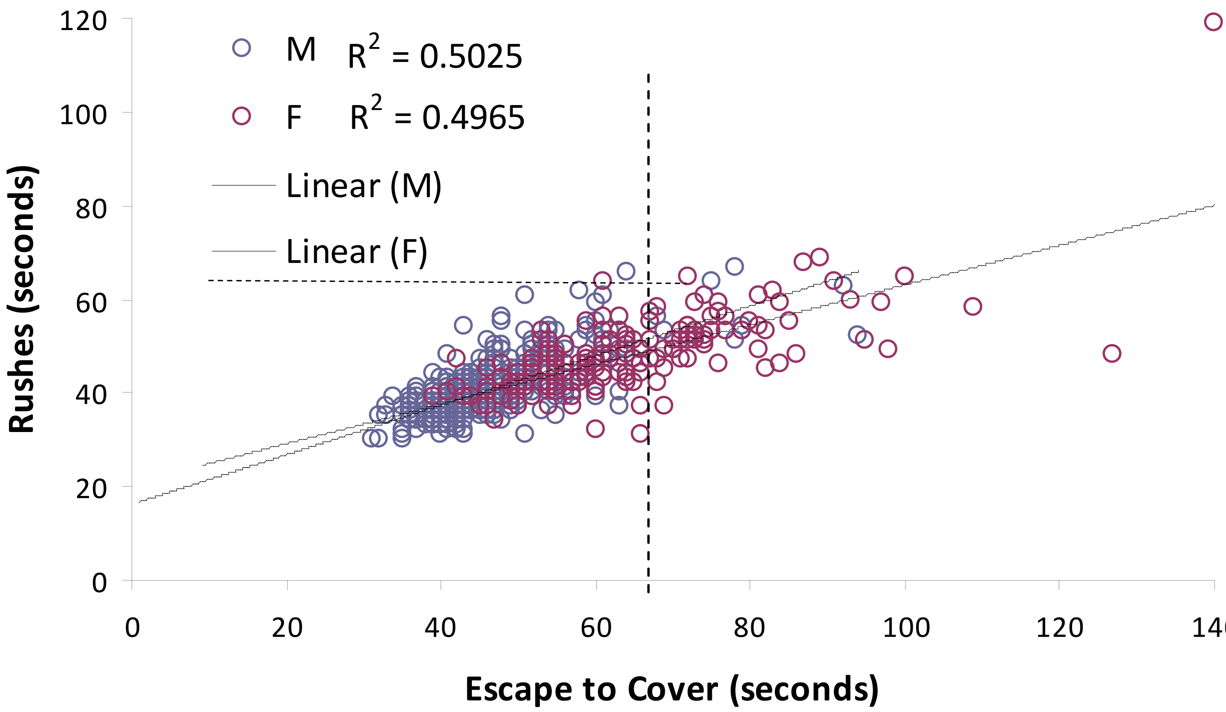


Figure 3-21: The Ability to Predict Escape to Cover from Rushes for Males and Females [53].

Where the same predictive test predicts performance on multiple task simulations, the most demanding standard should be chosen to ensure all task simulations can be achieved. For example, for the CAF the *Sandbag Lift* (a predictive test) was the strongest predictor of both tasks *Sandbag Fortification* and *Picking and Digging*. The simple regressions yielded a *Sandbag Lift* time of 3:30 min:sec for successful performance on *Picking and Digging* whereas the *Sandbag Lift* standard was 5:50 min:sec when related to *Sandbag Fortification*. In this case, a time standard of 3:30 min:sec was set for the *Sandbag Lift* test component.

As previously mentioned, the strength in a regression can decrease as the dependent variable score moves further from the mean value of the chosen dependant variable. One method to consider this effect is described by Beck et al., [23] as the K-Fold analysis. In this method you hold-out a part of the available dataset from the whole learning process (recommended 5 – 10 x) and you then train your model using the remaining portion of your dataset.

The part of the dataset you held-out during training is now used as the test set, since it's actually unseen from the model's point of view. The testing framework will usually give an indication of the model's effectiveness, e.g., the mean squared error in case of linear regression. If the indicator in the previous step suggests that the model works well enough, it is then trained using the entire dataset. The K-fold Cross-Validation is a powerful practical tool to quantify a model's performance and has been shown to be accurate [23].

3.6.1 Adverse Impact and Predictive Bias

When PES are challenged it is most likely on the basis that the PES is perceived as discriminatory. Any test, by nature, is designed to be discriminatory. The Bona Fide Occupational Requirements and Bona Fide Justifications document [8] state that any investigation into a complaint against a PES may include any of the following issues:

- Does the standard exclude members of a particular group based on impressionistic assumptions?
- Does the standard treat some more harshly than others?
- Were alternative standards considered?
  - If so, why weren't the alternatives implemented and why was this particular standard chosen instead of others?
- Is the standard the least discriminatory means of accomplishing the purpose?
- Is it necessary that all employees meet a single standard, or could varying standards be adopted?

There are multiple dimensions to fairness [55], but in practice it "is not simply a matter of whether or not test score averages differ by...[group], but whether or not there are differences in test score predictions by group" [47]. If the predictions are equivalent (i.e., no differences in [estimated relationships between test scores and performance measures]), then there is no bias" [44]. This is an important concept to understand and apply to PES development.

3.6.1.1 Adverse Impact

There are few quantitative methods to determine if bias exists in a test, calculate Adverse Impact (AI), Standard Deviation (Z test), or Fishers Exact Test. However just because AI is not present, predictive bias which can occur through the use of predictive tests must be considered for population subgroups such as females. These tests will be described in this section.

To assess if a PES will exclude members of a particular group the assessment of "adverse impact" is often made. The current calculation applied to assess Adverse Impact originates from Title VII of the Civil Rights Act (CRA) of 1964, the Age, Discrimination in Employment Act (ADEA) of 1967, and Americans with Disabilities Act (ADA) of 1990. Title VII prohibits employment discrimination because of "race, color, religion, sex, and national origin'' by employers, labor organisations, and employment agencies (EEOC, 1978). Adverse impact can occur in the hiring, promotion, training, and transfer processes. In the context of physical performance, adverse impact commonly occurs in the applicant selection setting. When adverse impact occurs in the physical performance domain, it can result in different passing rates on a test between men and women. However, adverse impact has been observed in relation to age, race, and national origin in physical assessments depending upon the applicant pool.

There are several methods to determine whether adverse impact exists. The three most common approaches include:

- 4/5ths or 80% rule;
- 2 standard deviation or Z test; and
- Fisher's Exact Test.

For purposes of this adverse impact explanation, an assessment (e.g., physical test) is used to demonstrate the methods for calculating adverse impact. Although not legally mandated, the Equal Employment Opportunities Commission (US) guidelines used the four-fifths (4/5ths) rule, i.e., if the minority group pass rate is less than 80%, of the pass rate of the majority group pass rate, adverse impact exists. Trial courts need not adhere to the 4/5ths rule, but legal history shows that the 4/5ths rule was viewed favorably by the courts [45].

In 2010, 70 experts in adverse impact from the United States attended a conference designed to identify the best technical practices on how to conduct adverse impact analysis [56]. These experts included labor economists, industrial-organisational psychologists, plaintiff and defence attorneys, human resources practitioners, and officials from the Equal Employment Opportunity Commission (EEOC) and the Office of Federal Contract Compliance Programs (OFCCP). Data from these experts showed the 4/5ths rule produced false positives too often and did not flag true differences [55]. Fifty-eight percent (58%) of the experts found statistical significance tests (e.g., Fisher's Exact Test) very useful in identifying adverse impact, while only 19.2% found the 4/5ths rule very useful. That being said, the 4/5ths rule is the first calculation taken to determine adverse impact and cited in the EEOC ***Uniform Guidelines*** **(1978)**. Note that significant mean differences between the majority and minority groups on a test battery does not necessarily indicate there is adverse impact.

Within the statistical methods (e.g., 2 standard deviation), there are multiple variations on the formulae used. In most instances they will result in the same conclusion. However, it's important to emphasize that multiple methods should be used when there are very large and very small sample sizes. Further, as the sample size increases, the likelihood of a statistically significant result increases. Thus, multiple approaches for calculating adverse impact should be used. Although other methods exist, this discussion focuses on the three commonly agreed upon methods listed above, for a one page review Figure 3-22, **Determining Adverse Impact** briefly describes each test that is explained fully here.

3.6.1.1.1 4/5 ths or 80% Rule

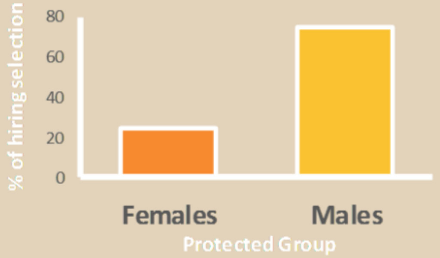
The 4/5ths rule is a comparison of the selection ratio or passing ratio on an assessment between the group with the highest select ratio or majority group (e.g., men, individuals under 40 years old) and groups with lower selection rates or minority groups (e.g., women, Blacks, individuals ≥ 40 years old). If the ratio between the majority and minority groups is less than 80%, there is adverse impact. However, large differences shown by the 4/5ths rule may not indicate adverse impact if the sample size(s) is small [57]. The steps below outline the calculation of the 4/5ths rule. Table 3-4 provides an example of the calculation.

Table 3-4: Example of 4/5ths or 80% Rule.

EEO Group	# of Applicants	# Passed Test	Percent Passed Test	Adverse Impact Ratio
Men	453	430	0.949	0.676
Women	95	61	0.642	

Determining Adverse Impact

(Gebhardt et al, 2018)



Adverse impact is defined as the presence of disparity in employment selection rates resulting in negative treatment of a person or groups such as: age, race, disability, religion, sex, sexual orientation, gender identity, and marital/family status. In the context of physical performance standards, adverse impact is typically found in the applicant selection process. This is commonly reflected by different passing rates on a test between men and women. Discrimination towards age, race, and national origin have also been observed in physical standards. Several methods to detect and avoid adverse impacts are outlined in the following section.

(Canadian Human Rights Act c H-6, 1985)

Methods

To demonstrate methods 1-3, the following dataset is used for sample calculations

Groups	# Completed Test	Total # Passed	Total # Failed	Pass Rate %	Expected # to pass
Male	135	125	10	93	115
Female	38	23	15	61	33
Total	173	148	25	86	148

Research show the 4/5ths rule to often produce false positives. 56% of experts found Fisher's exact test to be effective, while 19.2% found 4/5th rule to be useful. It is suggested that **MULTIPLE** methods be used for very large or small samples in detecting **adverse impact.**

(Cohen et al., 2010)

1) 4/5ths or 80% Rule

Comparing passing ratios on assessment between highest ratio majority group with lower selection ratio minority groups. If ratio between minority/majority passing rate is less than 80%, there is **adverse impact.**

%Pass minority group/ %Pass majority group< 0.80

Ex. 61% / 93% = 0.65

0.65 < 0.80

Test suggests adverse impact

2) 2 Standard Deviation or Z test

Investigates whether a difference in passing rate is due to chance at a probability valve of 0.05. A difference of 2 standard deviations indicates **adverse impact** when comparing the expected number of passes to the actual number .

$$SD = \times \sqrt{\frac{(\text{\# of f. apps})}{(\text{\# of tot. apps})} \times \frac{(\text{\# of m. apps})}{(\text{\# of tot. apps})} \times (\text{\# of tot. apps selected})}$$

Ex. SD = 5, 2SD = 10

Female expected = 33

$33_{expected} - 10_{2SD} = 23$

Actual # of pass within 2SD, **NO adverse impact**

3) Fischer's Exact Test

A statistical test, obtaining the possibility of all combinations, to determine whether the difference in passing rate is due to chance . **Adverse impact** is detected with a significant probability value at 0.05. Two -tailed test shown to be more accurate, however one tail test should be used when minority group consistently passes with lower rate (hypothesizing disparity

Fischer's Equation

$$p = \frac{((a+b)!\,(c+d)!\,(a+c)!\,(b+d)!)}{(a!\,b!\,c!\,d!\,N!)}$$

a,b,c,d = frequency of pass/fail for male/female (2x2 table)

N = total sample size

p= 0.00001, <0.05

Test suggests adverse impact

NATO HFM RTG 269 Combat Integration: Implications for Physical Employment Standards

Figure 3-22: Summary of the three Methods to Evaluate Adverse Impact According to Gebhardt [58].

Calculate the rate of selection for the majority and minority groups by dividing the number of individuals who passed the assessment by the total number of individuals in the majority and minority group who took the assessment.

Determine which group has the highest passing rate. This is typically men for physical testing.

Calculate the adverse impact ratio by dividing the passing rate for the minority group by the majority group's passing rate.

Determine whether this ratio is equal to or greater than 80%. If it is less than 80% or 0.80, there is adverse impact.

Table 3-4 shows that the adverse impact ratio was 0.676 or 67.6%, which is less than the 80% required. Thus, there was adverse impact on women.

3.6.1.1.2 2 Standard Deviation or Z Test

The 2 Standard Deviation (SD) or Z test investigates whether the difference in passing rates is due to chance and uses a probability value of 0.05 that indicates the difference was not due to chance. The 2 SD method compares the expected passing rates to actual passing rates. A difference two (2.00) or more standard deviations in passing rates indicates adverse impact is present.

Table 3-5 contains the applicant numbers used to calculate whether there was adverse impact with the test. There were 173 applicants who completed the test and a total of 148 passed the test. Thus, the expected pass rate for men and women is 85.55%. To determine whether the total number of women passing was within 2 standard deviations, the standard deviation and the passing rates are used.

Table 3-5: Example of 2 Standard Deviation / Z Test.

EEO Group	# Completed Test	Total Passed Test	Overall Pass Rate	Expected # to Pass
Men	135	125		115.5
Women	38	23		32.5
Total	173	148	85.55%	

Shown below is the formula for calculation of the standard deviation using the number of male and female applicants and the total number of applicants who passed the test. Note that *min apps* stands for the minority group applicants (e.g., women) and *non-min apps* stands for majority group applicants (e.g., men). The *No. of apps selected* is the total number of applicants who passed the test.

$$SD = \sqrt{\frac{(No. of \min apps)}{(No. of\ total\ apps)} x \frac{(No. of\ non - \min apps)}{(No. of\ total\ apps)}\ x\ (No\ of\ apps\ selected)} \quad (3\text{-}5)$$

The steps below outline calculation of to determine if the women's passing rate results in adverse impact.

Calculate the standard deviation value using the above formula:

√ {(38/173) * (135/173) * 148}
√ {(0.2197) * (0.7803) * 148}
√ 25.3719
= 5.04

Calculate the expected passing rate for the minority group (women) based on the passing rate for the total sample.

Passing rate for total sample = 85.55%

Expected passing rate for minority group (women) = 32.5

Compare number of women actually passing the test (23) to the expected number to pass the test (32.5) minus two (2) standard deviations (5.04 * 2).

32.5 – (2*5.04) = 22.42

There was no adverse impact because the number of women who passed the test (23) exceeded the predicted number of women passing the test minus 2 standard deviations (22.41).

3.6.1.1.3 Fisher's Exact Test

Fisher's Exact Test uses a two-tail statistical test to determine whether the difference in pass rates is beyond chance. If the probability value is significant at the 0.05 level, there is adverse impact. This test uses the formula below to obtain the probability of all possible combinations. One of the issues with the Fisher's Exact Test is whether to use a one- or two-tailed test. A survey of 70+ experts in adverse impact (statisticians, labor economists, I/O psychologists) showed a two-tailed test to be more accurate and one must consider sample sizes [59]. However, Tippins [57] stated that a one-tailed test should be used when the minority group always passes at a lower rate similar to hypothesizing a specific disparity direction. In physical testing, the minority group (women) do have a lower passing rate than men almost all of the time.

Shown below is the formula for the Fisher's Exact Test. Since it contains factorials (e.g., 4! or 1*2*3*4 = 24), one typically uses a statistical calculator or program to generate the value and significance level. Fisher's Exact Test is:

$$p = \frac{(a+b)!\,(c+d)!\,(a+c)!\,(b+d)!}{3.4.2\quad (a!\,b!\,c!\,d!\,N!)} \tag{3-6}$$

where a, b, c and d are the individual frequencies in a 2x2 table. "N" is the total frequency (N=173). Calculations using this formula resulted in a value of 0.00001 using a 2-tailed test, thus significant at $p<.05$ and indicating adverse impact is present. Notice that the Fisher's Exact Test result showed adverse impact while the 2 standard deviation test found no adverse impact.

These three tests are depicted and summarized in Figure 3-22. Use of a specific approach depends on the sample size of the groups being compared, applicant flow rates, and how the employer makes hiring decisions. As stated previously, use of more than one test is appropriate when addressing adverse impact. Finally, it should be noted that a finding of adverse impact does not necessarily mean the test is discriminatory. If a large sample of applicants take the test, it is likely to find adverse impact. However, if the test is job-related and supported by strong validity evidence and business necessity, the adverse impact can be acceptable.

3.6.2 Predictive Bias

Reilly et al. [6] established 3 criteria for the design of military PES to apply in order to avoid bias by design:

- Diversify the SME to ensure representation of all minority groups;
- Diversify the participants used during field measurements of the physiological demands; and
- Be cognizant of the bias inherent in predictive tests = Predictive Bias.

Predictive bias can take two forms. First, it can occur when predictive validity differs by group, a phenomenon known as differential validity. If the test is a better predictor of performance for one group than it is for another, then the test is considered biased against the group with the lower predictive validity [44]. Second, it can occur when the predictive validity is equivalent for both groups but the test still under predicts one group's performance relative to another group. For example, if, for men, a score of 10 on a strength test suggests that they will fail and the same test is used for both men and women, then a 10 for women should have the same expected outcome – namely, failure. If, however, a study shows that a score of 10 would predict that women would, on average, succeed on the job when men with the same score would, on average, fail, the test would be under-predicting female performance. Both types of bias need to be examined. If a test is discovered to exhibit either type of bias, it should not be used [44]. For example, if using abdominal circumference to identify risk of cardiovascular disease, short subjects have higher levels of risk factors and a 30% higher prevalence of the metabolic syndrome than tall subjects if grouped by abdominal circumference but not if grouped by weight-to-height ratio. (N = 6971) [60].

3.6.3 Accommodation

Adverse Impact has been justified successfully if individuals with equal probabilities of success have equal probabilities of being hired. If a group has the *potential* to pass the physical fitness test, it is questionable if the PES is considered to have adverse impact. A good example of this is the difference between a Fitness Standard PES, and a Height Standard PES. If the fitness requirements of the test are attainable (through training) by the population and are reasonable, the potential is not different for males vs. females. However, if there is a height standard not related to the job demands, which is greater than 80% of females, this cannot be achieved by the females (for example with physical training). (Case Study, see Figure 3-23.)

PES CASE LAW: CANADA

Chapdelaine v. Air Canada (1991)

The Saskatchewan Government and General Employees Union (2015)

Two cases which have illustrated this are *Chapdelaine v. Air Canada, 1991 CanLII 553 (CHRT),* where Air Canada's height policy, although perhaps "on its face neutral" in its application, operated to deprive 82% of all Canadian women and only 11% of all Canadian men between the ages of 20 and 29 from the opportunity for employment as a pilot. Therefore, the court ruling was such that considerably more women than men were adversely affected by Air Canada's height policy.

On the contrary in *The Government of Saskatchewan Ministry of Environment vs The Saskatchewan Government and General Employees Union (2015),* The Saskatchewan Wildland Firefighters grieved the newly implemented PES (the WFX-Fit test) on the basis that it was discriminatory against older employees and females.

Interestingly, this test was specifically designed with regard to the implementation of the time standard, to avoid Adverse Impact. The evidence explained to the court how the passing time for the test was chosen using statistics based on fire fighter performance results from incumbents across Canada. The fitness test developer intentionally chosen the passing time at which 82.5% of the women who participated in the development process had completed the test. However, this passing time was not linked to a BFOR, and therefore it was found to be arbitrary. There was no evidence presented in this case that those who may fall within the 17.5% group were incapable of performing the work of a wildland fire fighter in a safe and efficient manner. The arbitrator found that since this kind of arbitrariness is the essence of discrimination, the fitness test was *prima facie* discriminatory.

Figure 3-23: PES Case Law: Canada – Chapdelaine v Air Canada (1991). The Saskatchewan Government and General Employees Union (2015).

There are many acceptable methods of providing accommodation to groups who are discriminated against, the most effective, in the case of PES, are physical training programs sponsored and supported by the employer. For example, Jamnick et al., [15] demonstrated that, males and females can improve completion time of a PES by 10 – 11% with familiarisation and an additional 18 – 22% with a customized six week physical fitness training program. The potential for predictive selections test and task simulation performance to improve with training will be discussed in Chapter 5.

3.6.4 Bias Mitigation

Bias mitigation has been discussed throughout this section and quickly reviewed here as the best method to ensure one has done his/her due diligence in PES design are:

- Know the inherent bias of your potential PST and design your test battery to avoid the use of test with bias if possible.
- Know the potential of your PST to improve with training.
- Understand the relationships between PST and task simulations and make every attempt to select your PST with the goal of combining male and female data.
- Examine the extreme ends of your regressions and determine if there is bias for one particular group at the level for which you will be setting the standard.
- Maximise physical training programming on the **Components of Fitness** for which the incumbents demonstrate they are lacking, based on PES results.
- Dictate form as little as possible as long as safety is considered as anthropometrics may mean there are more or less efficient ways to complete a task to the standard.

The safety risk of providing accommodation that lowers the standards for unsuccessful applicants or ‘‘unsafe and inefficient’’ incumbent workers would also constitute ‘‘undue hardship’’ for the employer [61]. This important resolution was legally enforced by a revision to the Criminal Code of Canada in 2004 (Bill C-45), which legislated that supervisors and management personnel must prevent ‘‘workplace negligence’’ or they could be deemed careless and liable if they fail to take reasonable measures. This is defined as ‘‘ensure the bodily safety of persons doing the work or task, and when public safety workers require physiological attributes necessary to avoid foreseeable risks, an employer has a duty of care or due diligence responsibility to ensure that those physiological attributes are present’’ [15].

In conclusion, to lower a standard which has been established as a BFOR, with the intention of avoiding Adverse Impact is not advised. This is not a method of accommodation. PES that have utility usually have adverse impact. Although adverse impact is undesirable, it is acceptable as long as the test is valid, job-related, and fair [45]. It is generally accepted that the more physically demanding the job, the more likely a test will fail the EEOC Guidelines 4/5ths rule.

((See Section 3.8 for further background reading.)

3.7 VALIDATION AND RELIABILITY

The most recent publication on validation strategies for PES currently is by Milligan et al. [16]. This paper discusses various methods of validation and highlights concerns such as reliability and validity theory detailing how they are interconnected. For example, reliability is integral to validity in that a selection test or PES cannot be considered valid if it is not reliable.

Validity evidence was addressed in this paper from four perspectives: 1) Content; 2) Logical; 3) Criterion; and 4) Construct; whilst reliability was discussed from the perspectives of systematic and random error.

The following definitions are given for the four types of validity and visualized in Figure 3-24:

Content validity is defined as the degree to which a test adequately samples what was covered by the critical tasks required by the job, determined by qualitative and quantitative measures. *Use of SME and Methods of Best Practice (Qualitative/Subjective). Collection of physical and physiological data (Quantitative/Objective)* [62].

Logical validity, more commonly known as "face" validity, is achieved when a task analysis includes consultation with subject matter experts (SME), experienced supervisors, and employees, and is most apparent when direct task simulations are used as selection tests in developing PES [62]. Evidence is based on subjective assessment of content (e.g., questionnaires), non-objective statistical evidence can be provided for logical validity. Minimum Occupational Performance Standard setting: *SME though bookmarking (Qualitative/Subjective), SETS (standards through scenarios with SME) (Qualitative/Subjective).*

Criterion Validity. There are two types of *criterion-related evidence*: concurrent and predictive.

Concurrent validity is usually employed when a criterion measure or gold standard test is to be substituted by a simple or easily administered alternative. For example, directly measured maximal oxygen uptake is considered the gold standard measure of aerobic power. Validating the use of indirect assessments of maximum oxygen uptake (e.g., shuttle runs or step tests) as a replacement for the direct laboratory assessment of maximal oxygen uptake is concurrent validity. This could also apply to the relationship between predictor tests and their simulations. *Collection of physical and physiological data (Quantitative/Objective). Consider reliability.*

Predictive validity; this is especially important for determining the predictive capability of a test, but this validity check is often neglected or not considered part of the research plan. Predictive validity is a measure of a test's ability to predict future outcomes [63]. Predictive validity is commonly used in pre-employment tests to determine how well a test can predict future success in the occupation. *(Quantitative/Objective).*

Construct validity determines whether a PST measures the same constructs as those that actually govern the physical performance of the critical task. It is made up of:

a) Convergent validity (i.e., constructs that theoretically should be related to each other are, in fact, observed to be related to each other); and

b) Discriminant validity (i.e., constructs that theoretically should not be related to each other are, in fact, observed to not be related to each other).

Collection of physical and physiological data (Quantitative/Objective). Consider reliability. A PES demonstrates construct validity if it accurately differentiates between those individuals who are, and are not, capable of performing a critical task to the Minimum Performance Standard (MPS).

3.7.1 Validating Task Simulations

Often when task simulations are developed, to be administered as a PES, they are modified in order to facilitate administration to a large population with minimal equipment, and even indoors to mitigate the effects of the environment outside, increasing their **Feasibility** (Figure 3-18). For the CAF the physiological demands of three of the six criterion tasks (critical and essential) were originally quantified outside in their field context. When these tasks were simulated inside and administered as tests, a validation study was necessary to determine how accurate these simulations were at eliciting the same physiological response. This is the assessment of Content Validity as the task is a replica of the original task, it already demonstrates logical validity. These three tasks were:

1) Transporting pickets and wire for fence construction
2) Escape to cover, and
3) Picking and digging.

This research required quantification of the metabolic demands of these tasks at the same pace (minimum standard) by direct measure with a portable metabolic analyzer.

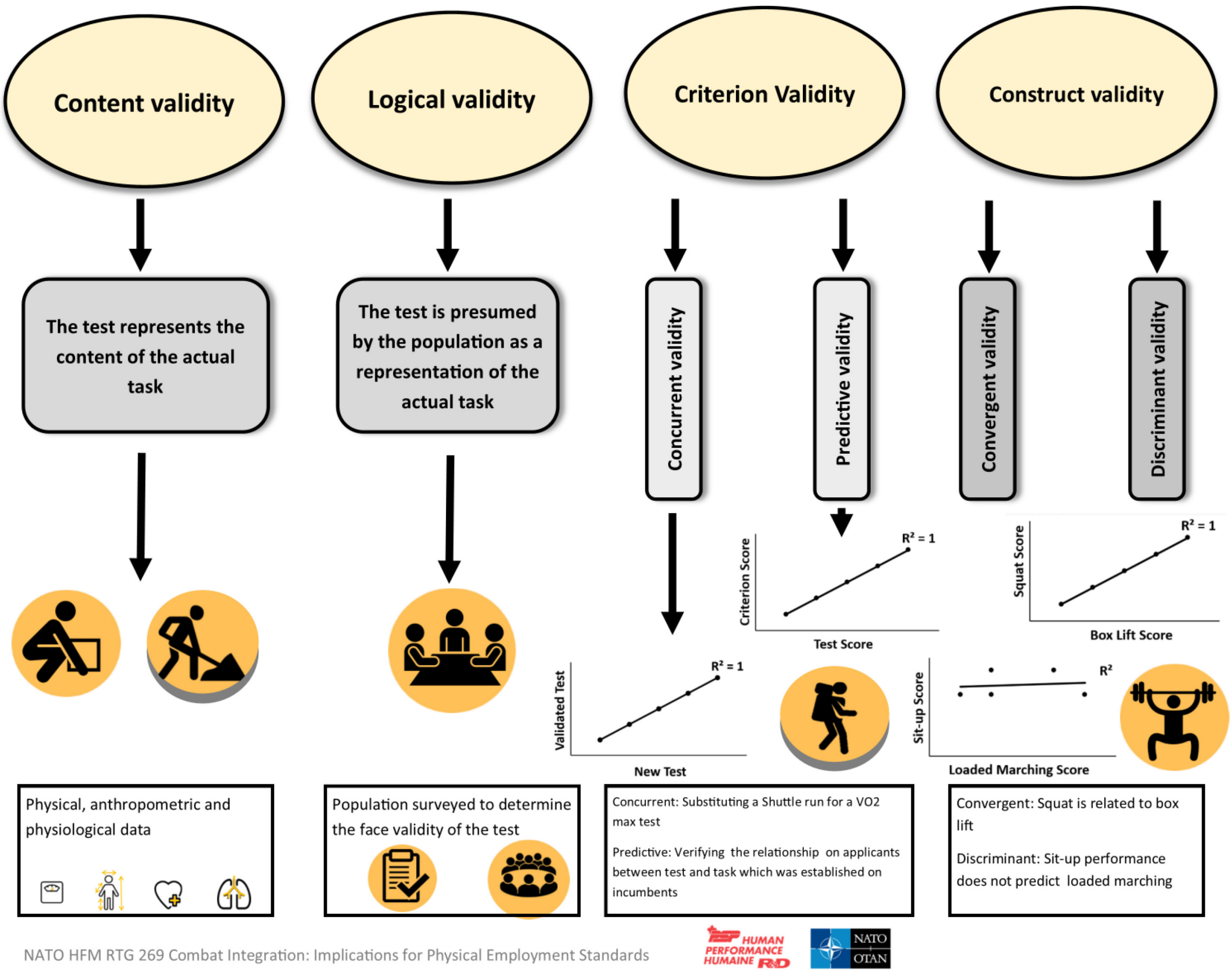


Figure 3-24: Flow Chart of Validity Concepts to Consider (Based on Concepts Presented in Milligan et al., [16]).

Transporting Pickets and Wire for Fence Construction. With regard to the physiological demands for carrying pickets and wire (paced at the minimum operational standard of 23 minutes), the oxygen consumption (VO_2) demonstrated a significant difference between the field and the indoor simulation. The VO_2 demands were significantly less (22 – 24% lower) on the gym surface than in the field ($P < 0.01$). The subsequent gym trials were then paced to meet the oxygen demand of the field trial. The resulting time to complete the gym simulation was, on average (SD) 17:05(1:26 s) (N = 16, 8F, 8M). Therefore, although the gold standard determined in the field was 23:00, the standard to achieve this task when performing it on a gym floor was adjusted to 17:05 to ensure the energy costs were similar to the field-based task.

Escape to Cover. Similar work was performed on the Escape to Cover task and results demonstrated that for this task, for which the minimum operational standard was much shorter (68 s) there was no significant difference in oxygen consumption when comparing the field trials to those performed in the gymnasium. Therefore, the standard remained at 68 seconds.

Picking and Digging. To validate the tasks of picking and digging was more complicated as this task could be performed on various field surfaces. Hard compact ground was identified as the operational requirement and the force required to disrupt the ground was measured. A 36” standard issue 6 lb pickaxe instrumented with a calibrated strain gauge was used to capture this force. A 6 lb sledgehammer with 36” handle was also instrumented with a calibrated strain gauge for the indoor field simulation. The task simulation consisted of a Keiser forcible entry machine and the sledgehammer was used to move a weighted beam on the Keiser, simulating picking. This Keiser allowed the safe simulation of the posture and movements of using the pickaxe. The strain gauge forces on both the pickaxe and the sledgehammer were matched to quantify the distance (4 m) the weighted beam must be moved to equate to the force required to disrupt 180 L of hard compact ground.

The digging component of the task was also validated by comparing the metabolic demand of digging in the hard compact ground measured in the field, after it had been broken up with a pickaxe. To ensure steady state, oxygen consumption was measured for 3 minutes continuously rather than the 60 s:30 s work/rest ratio of the task. The pace and volume of the substrate dug during the 3 minutes was recorded and repeated using ¾” – 1¼” river stones. River stones were chosen for both a scientific and a logistical perspective. The size of the stones is standardised and prevent full penetration of the shovel blade therefore better simulating the task where only the top layer of hard compact ground that has been loosened by the pickaxe can be excavated. The metabolic cost of digging the same volume at a set pace in the two substrates was very similar and considering the variability in substrate and how this would affect the oxygen demands, a force analysis was chosen as the most reliable way to validate the simulation. (See Figure 3-25.)

Figure 3-25: Simulations of Picking and Digging for the Canadian Armed Forces.

3.7.2 Validating Predictive Tests

If a predictive test is chosen, the score to achieve on the predictive test is designed to determine competence on the field tasks, and this relationship may require Criterion Validity (concurrent and predictive) and Construct Validity. The regression or relationship between the predictive test and the field tests would have been measured on a cohort population and therefore a subset cohort should be tested to determine if this relationship maintains its predictive strength, and its predictive validity.

US Army Active duty Soldiers were used in all three phases of their PES research process. At the time of the research, trained Soldiers performed both the OPAT (Occupational Physical Assessment Test) and the Critical/Essential criterion tasks within a two week period (concurrent validation). However, the intent of the PES (OPAT) was to test new recruits for entrance into the Army [64]. Therefore, in order to ensure the OPAT would correctly identify new recruits with the potential to perform the physically demanding tasks of their jobs, an additional study was conducted to determine the predictive validity of the OPAT in new recruits. 1,181 recruits (948 men, 233 women) completed the OPAT within the first two weeks of starting Initial Entry Training (IET) and 741 (608 men, 133 women) returned to perform the Critical/Essential criterion tasks (Criterion Measure Task Simulations) within five weeks of completion of IET [65].

These simulations included a tactical foot movement, casualty drag, casualty evacuation from a vehicle turret, sandbag carry to build a fighting position, moving under direct fire, loading a Field Artillery ammunition supply vehicle, reloading a tank, and loading the main gun on a tank. Regression analyses were conducted to predict performance on the Critical/Essential criterion tasks (criterion measure task simulations [CMTS]) from the OPAT scores. The four event OPAT test battery correctly identified 76% of new recruits who were able to perform the physically demanding tasks of their assigned combat arms MOSs by the end of IET, accounting for over 62% of the variance in CMTS performance. Therefore, both Criterion (predictive) and Construct Validity were quantified and evidenced by the results [65].

3.7.3 Reliability of Predictive and Task Based Tests

Reliability refers to a measure of consistency (reproducibility) within the data. To ensure minimal measurement error or optimal reliability, the systematic (learning or fatigue) and the random error (biological or mechanical variation of applicants and/or incumbents undertaking the tests or those administering the tests) must be considered and assessed [16], [29], [32], [65], [66]. It has been suggested that biological variation should form a large proportion of random error with the responsibility falling to the test administrators to minimize technical and environmental variability (e.g., equipment calibration, test circuit set-up) [29]. The number of trials required to achieve reliable results will depend on the number of degrees of freedom in the test and potential for improvement through learning or strategy development [65]. Previous research on skill based, loaded, circuit-type fitness tests indicate that maximum performance is achieved by the fourth trial [67]. This is further supported by Pandorf et al. [31], who evaluated the reliability of a six element, indoor, obstacle course designed to simulate impediments to movement of a soldier during a conflict.

There are various ways of evaluating reliability and variability of performance over time including the Coefficient of Variation (CV); Typical Error (TE) (the amount of inherent test error that could be reasonably expected on any given day in which the individual completes a “best effort” trial on a given test item) [68]; Smallest Worthwhile Change (SWC) as a measure of change between repeated exposures [68]; Minimal Detectable Change (MDC) (change in which systematic bias and random error are accounted for with a 95% confidence limit) [66], [68], [69]; and intra-class correlation coefficient (ICC) [29], also known as the reliability coefficient [68]. Research indicates that for a test to be reliable, the typical error should not be greater than the SWC [68], ICC should be 0.75 – 1.00 [70] and CV lower than 5% (ideally closer to 2.2%) for human performance testing [29]. Reliability studies should be performed with a “representative cohort of individuals”, a sample from the wider population of those that could apply for a job as well as those that are currently in the job [13], [16].

Pandorf’s research on the six-element indoor obstacle course identified performance improvements of 4% from trials 1 – 2 and 3% from trials 2 – 3, and an intra-class correlation coefficient (ICC) of 0.92 on trial 4 as well as a mean Coefficient of Variation (CV) of 4.1% across all trials. The results of a reliability study on the Canadian Armed Forces PES, the four component FORCE Evaluation, demonstrate that candidates should be provided with at least one retest if they have recently completed at least two previous “best effort” attempts as per the protocol [32]. In addition, the minimal detectable difference is given for each of the four components in seconds which identifies the threshold for subsequent action, either retest or remedial training, for those unable to meet the minimum standard [32].

A more complex test, performed in full fighting order (25 kgs) and as a circuit containing many more degrees of freedom: the Canadian Army Fitness Objective FORCE COMBAT™ reliability study identified significant changes in mean values between trials 1 – 2 and 2 – 3 and non-significant changes in trials 3 – 4 suggests that two practice trials are required before stable performance is achieved [65]. (See Section 3.8 for further background reading.)

3.8 BACKGROUND READING

Hanser, L.M., and Oguz, M. (2015). United States Service Academy Admissions: Selecting for Success at the Military Academy/West Point and as an Officer. Rand National Defense Research Institute Santa Monica CA.

Davis G. and Fernandes A. (1995). Physical Selection for Rescue Craft Crew, Offshore Technology Report. OTH 93408. Prepared for the Health and Safety Executive, University of Surrey.

Sharp, M.A., Foulis, S.A., Redmond, J.E., Canino, M.C., Cohen, B.S., Hauret, K., Frykman, P.N., Pierce, J.R., Westrick, R.B., Pacheco, B.M., Gebhardt, D.L., and Zambraski, E.J. (2018). Longitudinal Validation of the Occupational Physical Assessment Test (OPAT). USARIEM Technical Report T18-05, U.S. Army Institute of Environmental Medicine, Natick, MA, US. DTIC: AD1057800.

3.9 REFERENCES

[1] Gledhill, N. and Bonneau, J. (2000, September). Objectives, Process and Consensus Summary of the National Forum on bona fide Occupational Requirements. In Proceedings of the Consensus Forum on Establishing bona fide Requirements for Physically Demanding Occupations, Eds. N. Gledhill, J. Bonneau, and A. Salmon. Toronto, Canada, York University, 1-6.

[2] Deakin, J., Smith, J., Pelot, R., and Weber, C. (2000). Methodological considerations in the development of physical maintenance standards. In Proceedings of the Consensus Forum on Establishing bona fide Requirements for Physically Demanding Occupations. Eds. N. Gledhill, J. Bonneau, and A. Salmon. Toronto, Canada, York University, 17-21.

[3] Rayson M., Holliman, D., and Belyavin, A. (1997). OPL Report. Validation of Physical Selection Standards for Recruits. Technical Report Ref PSS(R)/Rep09.07/v1.

[4] Taylor, N.A., and Groeller, H. (2003). Work-Based Physiological Assessment of Physically-Demanding Trades: a Methodological Overview. Journal of Physiological Anthropology and Applied Human Science, 22(2), 73-81.

[5] Tipton, M.J., Milligan, G.S., and Reilly, T.J. (2013). Physiological Employment Standards I. Occupational Fitness Standards: Objectively Subjective?. European Journal of Applied Physiology, 113(10), 2435-2446.

[6] Reilly, T., Gebhardt, D., Billing, D., Greeves, J., and Sharp, M. (2015). Development and Implementation of Evidence-based Physical Employment Standards-Key Challenges in the Military Context 3rd ICSPP JSCR Supplement Paper, 29(11S) S28-S33.

[7] Payne, W., and Harvey, J. (2010). A Framework for the Design and Development of Physical Employment Tests and Standards. Ergonomics, 53(7), 858-871.

[8] Canadian Human Rights Act (R.S.C., 1985, c. H-6), Act current to 2019-05-07 and last amended on 2017-06-19. https://laws-lois.justice.gc.ca/eng/acts/h-6/fulltext.html.

[9] United Kingdom Equality Act 2010 (c15). https://www.legislation.gov.uk/ukpga/2010/15/contents.

[10] Civil Rights Employment Law USA Title VII of the Civil Rights Act of 1964.

[11] Act Relating To Equality And A Prohibition Against Discrimination (Equality and Anti-Discrimination Act), Ministry of Culture, Last consolidated LOV-2017-12-19-115 from 01.01.2018, Last update 03.06.2019, https://lovdata.no/dokument/NLE/lov/2017-06-16-51.

[12] Equality Act 2010 (UK), c 15.The Discrimination Act 1992, Australia (Cth).

[13] Equal Treatment Act 1994 (Algemene Wet Gelijke Behandeling). Adopted on 2 March 1994. https://www.equalrightstrust.org/content/netherlands-equal-treatment-act-1994.

[14] Lee-Bates, B., Billing, D.C., Caputi, P., Carstairs, G.L., Linnane, D., and Middleton, K. (2017). The Application of Subjective Job Task Analysis Techniques in Physically Demanding Occupations: Evidence for the Presence of Self-Serving Bias. Ergonomics, 60(9), 1240-1249.

[15] Jamnik, V., Gumienak, R., and Gledhill, N. (2013). Developing legally Defensible Physiological Employment Standards for Prominent Physically Demanding Public Safety Occupations: A Canadian Perspective. European Journal of Applied Physiology, 113(10), 2447-2457.

[16] Milligan, G.S., Reilly, T.J., Zumbo, B.D., and Tipton, M.J. (2016). Validity and Reliability of Physical Employment Standards. Applied Physiology, Nutrition, and Metabolism, 41(6), S83-S91.

[17] Reilly, T. and Tipton M. (2005). Fitness Standards for Marine Rescue Operations Around the British Isles. Environmental Ergonomics XI. Proceedings of the 11th International Conference: 22-26 May, Ystad, Sweden.

[18] Larsen, B., and Aisbett, B. (2012). Subjective Job Task Analyses for Physically Demanding Occupations: What is Best Practice?. Ergonomics, 55(10), 1266-1277.

[19] Bowling, A. (2005). Mode of Questionnaire Administration Can Have Serious Effects on Data Quality. Journal of Public Health, 27(3), 281-291.

[20] Taylor, N.A., Fullagar, H.H., Sampson, J.A., Notley, S.R., Burley, S.D., Lee, D.S., and Groeller, H. (2015). Employment Standards for Australian Urban Firefighters: Part 2 The Physiological Demands and the Criterion Tasks. Journal of Occupational and Environmental Medicine, 57(10), 1072-1082.

[21] Brown P.E.H., Gunner, F., and Fallowfield, J. (2018). The Identification of Measurement Error from Responses to Control Questions Embedded in an Online Survey. 3rd Conference on Physical Employment Standards. ISBN:978-1-5272-2601-2: p. 66.

[22] Jamnik, V.K., Thomas, S.G., Shaw, J.A., and Gledhill, N. (2010). Identification and Characterization of the Critical Physically Demanding Tasks Encountered by Correctional Officers. Applied Physiology, Nutrition, and Metabolism, 35(1), 45-58.

[23] Beck, B., Billing, D.C., and Carr, A.J. (2016). Developing Physical and Physiological Employment Standards: Translation of Job Analysis Findings to Assessments and Performance Standards – A Systematic Review. International Journal of Industrial Ergonomics, 56, 9-16.

[24] Middleton, K.J., and Carr, A.J. (2017). The Identification of Combat Survivability Tasks Associated with Naval Vessel Damage in Maritime Environments. Applied Ergonomics, 59, 27-33.

[25] Blacklock, R.E., Reilly, T.J., Spivock, M., Newton, P.S., and Olinek, S.M. (2015). Standard Establishment Through Scenarios (SETS): a New Technique for Occupational Fitness Standards. Work, 52(2), 375-383.

[26] Reilly, T., Wooler, A., and Tipton, M. (2005). Occupational Fitness Standards for Beach Lifeguards. Phase 1: the physiological Demands of Beach Lifeguarding. Occupational Medicine, 56(1), 6-11.

[27] Reilly, T., Iggleden, C., Gennser, M., and Tipton, M. (2005). Occupational Fitness Standards for Beach Lifeguards. Phase 2: The Development of an Easily Administered Fitness Test. Occupational Medicine, 56(1), 12-17.

[28] Case Study: Barr and Flannery vs. Treasury Board: Department of National Defense, (2006). https://unde-uedn.com/wp-content/uploads/2019/02/FR-Fitness-test-Barr-Fla_001.pdf.

[29] Boyd, L., Rogers, T., Docherty, D., and Petersen, S. (2014). Variability in Performance on a Work Simulation Test of Physical Fitness for Firefighters. Applied Physiology, Nutrition, and Metabolism, 40(4), 364-370.

[30] Pandorf, C.E., Nindl, B.C., Montain, S.J., Castellani, J.W., Frykman, P.N., Leone, C.D., and Harman, E.A. (2003). Reliability Assessment of Two Militarily Relevant Occupational Physical Performance Tests. Canadian Journal of Applied Physiology, 28(1), 27-37.

[31] Foulis, S.A., Redmond, J.E., Frykman, P.N., Warr, B.J., Zambraski, E.J., and Sharp, M.A. (2017). US Army Physical Demands Study: Reliability of Simulations of Physically Demanding Tasks Performed by Combat Arms Soldiers. The Journal of Strength & Conditioning Research, 31(12), 3245-3252.

[32] Stockbrugger, B.G., Reilly, T.J., Blacklock, R.E., and Gagnon, P.J. (2018). Reliability of the Individual Components of the Canadian Armed Forces Physical Employment Standard. Applied Physiology, Nutrition, and Metabolism, 43(7), 663-668.

[33] Sharp, M., Rosenberger, M., and Knapik, J. (2006). Common Military Task: Materials Handling (No. USARIEM/MN-MISC-06-31). Army Research Institute of Environmental Medicine, Military Nutrition DIV, Natick, MA, US.

[34] Milligan, G. (2013). Fitness Standards for the Maritime and Coastguard Agency and the Oil and Gas Industry (Doctoral dissertation, University of Portsmouth).

[35] Fass, B. (2017). Reducing EMS Provider Lift Injuries, Journal of Emergency Medical Services, 2017, Accessed on 25.01.2019, https://www.jems.com/articles/print/volume-42/issue-10/features/reducing-ems-provider-lift-injuries.html.

[36] Vanderburgh, P.M. (2008). Occupational Relevance and Body Mass Bias in Military Physical Fitness Tests. Medicine and Science in Sports and Exercise, 40(8), 1538-1545.

[37] Kirknes, J., Aandstad, A., and Stornæs, A.V. (2014). Innstilling – Revid ert fysisk testordning for Forsvaret. Forsvarets høgskole, Norges idrettshøgskole Forsvarets institutt.

[38] Whitehead, P.N., Schilling, B.K., Peterson, D.D., and Weiss, L.W. (2012). Possible New Modalities for the Navy Physical Readiness Test. Military Medicine, 177(11), 1417-1425.

[39] Loturco, I., D'Angelo, R.A., Fernandes, V., Gil, S., Kobal, R., Abad, C.C.C., ... and Nakamura, F.Y. (2015). Relationship Between Sprint Ability and Loaded/unloaded Jump Tests in Elite Sprinters. The Journal of Strength & Conditioning Research, 29(3), 758-764.

[40] NATO (1986). Research Study Group on Physical Fitness. Final report. Physical Fitness in Armed Forces. 1986.

[41] Hackett, D.A., Davies, T.B., Ibel, D., Cobley, S., and Sanders, R. (2018). Predictive Ability of the Medicine Ball Chest Throw and Vertical Jump Tests for Determining Muscular Strength and Power in Adolescents. Measurement in Physical Education and Exercise Science, 22(1), 79-87.

[42] Aandstad, A. and Kirknes, J. (2018). Validity and Reliability of Muscle Strength Field Tests for Use in the Norwegian Armed Forces. 3rd International Conference on Physical Employment Standards, Portsmouth, UK, 2018.

[43] Knapik, J.J., Jones, B.H., Sharp, M.A., Darakjy, S., and Jones, S. (2004). The Case for Pre-Enlistment Physical Fitness Testing: Research and Recommendations (No. USACHPPM-12-HF-01Q9D-04). Army Center for Health Promotion and Preventive Medicine Aberdeen Proving Ground MD.

[44] Hardison, C.M., Hosek, S.D., and Bird, Chloe, E. (2015). Defining Physical Standards for Physically Demanding Jobs: A Review of Methods. 2752-2759.

[45] Constable, S., and Palmer, B. (2000). The Process of Physical Fitness Standards Development (No. HSIAC-SOAR-2000-001). Human Systems Information Analysis Center Wright-Patterson AFB OH HSIAC Program Office.

[46] Courtright S.H., McCormick, B.W., Postlethwaite, B.E., Reeves, C.J., and Mount, M.K. (2013). A Meta-Analysis of Sex Differences in Physical Ability: Revised Estimates and Strategies for Reducing Differences in Selection Contexts. Journal of Applied Psychology; 98: 623-641.

[47] Gebhardt, D.L., and Baker, T.A. (2010). Physical Performance Tests. Handbook of Employee Selection, 277-298.

[48] Moriguchi, C.S., Carnaz, L., Júnior, L.C.M., Marklin, R.W., and Coury, H.J.C.G. (2012). Are Posture Data from Simulated Tasks Representative of Field Conditions? Case Study for Overhead Electric Utility Workers. Ergonomics, 55(11), 1382-1394.

[49] Rayson, M., Holliman, D., and Belyavin, A. (2000). Development of Physical Selection Procedures for the British Army. Phase 2: Relationship Between Physical Performance Tests and Criterion Tasks. Ergonomics, 43(1), 73-105.

[50] Ludbrook, J. (1997). Special Article Comparing Methods of Measurement. Clinical and Experimental Pharmacology and Physiology, 24(2), 193-203.

[51] Ludbrook, J. (2002). Statistical Techniques for Comparing Measurers and Methods of Measurement: A Critical Review. Clinical and Experimental Pharmacology and Physiology, 29(7), 527-536.

[52] Wilkinson, D.M., Blacker, S.D., Richmond, V.L., Rayson, M.P., and Bilzon, J.L. (2014). Relationship Between the 2.4-km Run and Multistage Shuttle Run Test Performance in Military Personnel. Military Medicine, 179(2), 203-207.

[53] Reilly, T. (2014). Gender Neutral to Gender Free: Physical Employment Standard Evolution in the Canadian Armed Forces. In International Congress on Soldiers Physical Performance. USAIREM, Boston, USA.

[54] Petersen, S.R., Anderson, G.S., Tipton, M.J., Docherty, D., Graham, T.E., Sharkey, B.J., and Taylor, N.A. (2016). Towards Best Practice in Physical and Physiological Employment Standards. Applied Physiology, Nutrition, and Metabolism, 41(6), S47-S62.

[55] American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational, & Psychological Testing (US). (1999). Standards for Educational and Psychological Testing. American Educational Research Association.

[56] Cohen, D.B., Aamodt, M.G., and Dunleavy, E.M. (2010). Technical Advisory Committee Report on Best Practices in Adverse Impact Analyses. Washington, DC: Center for Corporate Equality.

[57] Tippins, N.T. (2010). Adverse Impact in Employee Selection Procedures from the Perspective of an Organizational Consultant. Adverse Impact: Implications for Organizational Staffing and High Stakes Selection, 201-225.

[58] Gebhardt, D. (2019). Personal Communication at Meeting 5/6 of HFM 269, Melbourne Australia, 2018.

[59] Zedeck, S. (2010). Chapter 1: Adverse Impact History and Evolution. In J. Outzz (Ed.). Adverse Impact: Implications for Organizational Staffing and High Stakes Selection. New York, NY: Routledge.

[60] Schneider, H.J., Klotsche, J., Silber, S., Stalla, G.K., and Within, H.U. (2011). Measuring Abdominal Obesity: Effects of Height on Distribution of Cardiometabolic Risk Factors Risk Using Waist Circumference and Waist-To-Height Ratio, Diabetes Care, Vol 34, no.1

[61] Tipton, M.J., Milligan, G.S., and Reilly, T.J. (2013). Physiological Employment Standards I. Occupational Fitness Standards: Objectively Subjective?. European Journal of Applied Physiology, 113(10), 2435-2446.

[62] Zumbo, B.D. (2016). Standard-Setting Methodology: Establishing Performance Standards and setting Cut-Scores to Assist Score Interpretation. Applied Physiology, Nutrition, and Metabolism, 41(6), S74-S82.

[63] Lin, W.L. and Yao, G. (2014). Predictive Validity. In A.C. Michalos (Ed), Encyclopedia of Quality of Life and Well-Being Research. Springer, Dordrecht.

[64] Foulis, S.A., Sharp, M.A., Redmond, J.E., Frykman, P.N., Warr, B.J., Gebhardt, D.L., ... and Zambraski, E.J. (2017). US Army Physical Demands Study: Development of the Occupational Physical Assessment Test for Combat Arms Soldiers. Journal of Science and Medicine in Sport, 20, S74-S78.

[65] Reilly, T., Walsh, E., and Stockbrugger, B. (2019). Reliability of FORCE COMBAT™: A Canadian Army Fitness Objective. Journal of Science and Medicine in Sport, 22(5), 591-595.

[66] Atkinson, G., and Nevill, A.M. (1998). Statistical Methods for Assessing Measurement Error (Reliability) in Variables Relevant to Sports Medicine. Sports Medicine, 26(4), 217-238.

[67] Todd Rogers, W., Docherty, D., and Petersen, S. (2014). Establishment of Performance Standards and a Cut-score for the Canadian Forces Firefighter Physical Fitness Maintenance Evaluation (FF PFME). Ergonomics, 57(11), 1750-1759.

[68] Hopkins, W.G. (2000). Measures of Reliability in Sports Medicine and Science. Sports Medicine, 30(1), 1-15.

[69] Hachana, Y., Chaabène, H., Nabli, M. A., Attia, A., Moualhi, J., Farhat, N., and Elloumi, M. (2013). Test-Retest Reliability, Criterion-Related Validity, and Minimal Detectable change of the Illinois Agility Test in Male Team Sport Athletes. The Journal of Strength & Conditioning Research, 27(10).

[70] Cicchetti, D.V. (1994). Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. Psychological Assessment, 6(4), 284.

Chapter 4 – INCENTIVIZATION

N. Baumgartner
US Air Force
UNITED STATES

D. Leyk
Central Institute of the Bundeswehr Medical Service
GERMANY

T.J. Reilly
Department of National Defence
CANADA

4.1 INTRODUCTION

The agreed aim of this chapter is to address potential incentivization methods that military services may use to motivate military service personnel to meet and exceed physical fitness standards and enhance the quantity and quality of physical training. Section 1 provides a brief summary of the extensive literature on incentives used in the workplace. Sections 2 and 3 address sex-fair and sex-free tests to which military services may apply incentives. Finally, Section 4 addresses military-specific limitations and opportunities for incentivizing members to meet or exceed physical fitness standards and includes a summary table of incentive methodology.

4.2 BONUS AND MALUS IN THE WORKPLACE FOCUSING ON THE BEHAVIORAL ASPECTS OF INCENTIVE DESIGN

A plethora of literature on incentives that primarily addresses wellness in the workplace is summarized in this first section by Germany, with additional input from the United States.

4.2.1 Bonus and Malus Systems, Incentives and Physical Performance

Human behavior often falls short in serving individual as well as collective interests. People fail to engage in healthy eating, regular exercise, frugal spending, and wise investments. Who would pay taxes if it was an entirely voluntary contribution? In most societies there are rewarding (bonus) and penalizing (malus) public systems aimed to guide human behavior. These public bonus/malus systems exist in numerous spheres of life. Consider some examples.

A Swiss farmer recently pushed for a plebiscite to implement a monetary incentive system for abstaining from horn mutilations. As cows with full horns are more work, horn-supportive farmers were supposed to get monetary compensation for their extra work, serving animal well-being [1]. Companies' loyalty programs represent a bonus system that we are all familiar with. Tax incentives for retirement savings represent a public bonus system [2], [3]. A non-public bonus/malus system can be found in the business world, in form of executive compensation. Car insurers have been using bonus and malus systems for decades, accident-free driving reduces the premium while accumulating collisions leads to a rise in premiums. Private health insurers have a similar system in many countries, whereby a lower costing insurant pays less in premiums than the chronically ill. Many health insurers include health and wellness-checks as a requirement for the lowest available premium [4], [5]. Moreover, public health insurers in Germany offer bonuses in form of fitness courses, monetary contributions to fitness courses of third-party vendors, health-related apps for smartphones, monetary contributions for not-covered treatments, and even monetary contributions for the purchase of a smart watch or fitness tracker, and other benefits, in exchange for bonus-points. The insurant can gather points through participation in certified fitness courses, wellness check-ups and vaccinations [6].

Not only do health insurance companies use bonus systems to promote healthy lifestyles. Over the last several decades, employers have realized their vested interest in their employees' health, well-being, and resilience [7]. Amid a globalized economy, demographic change, and widespread unhealthy lifestyles, corporate success is affected by the health status and resilience of its employees. However, while health and wellness programs have become increasingly popular among employers [4], Leyk et al. report final participation rates of less than 18% in their case study of an agency of the Federal Armed Forces in Germany [8]. Other case studies have reported a positive return on investment [9], [10], [11]. Nevertheless, ongoing participation in health and wellness programs as well as long-lasting health-related benefits are not easily accomplished.

Even in hazardous professions, such as firefighter, overweight and prevalence of other risk factors are on the rise [12], [13]. As a consequence, in some local instances, health promotion programs tailored to the firefighter profession have been developed [14], [15]. Moreover, empirical evidence suggests a trend of decreasing physical fitness among soldiers [15], [16], [17], [18], [19]; see also the 2012 Technical Report NATO HFM RTG-178, "Impact of Lifestyle and Health Status on Military Fitness". In summary, there is a plethora of bonus/malus systems. There is also ample evidence for an overall decreasing physical fitness. Is there empirical evidence that physical fitness can be improved with incentives? We conducted a selective literature search to examine whether efficacy of health and fitness initiatives can be increased through reward systems.

4.2.2 Method

A search on October 26th 2018 was conducted on: https://www.ncbi.nlm.nih.gov/pubmed/ with the following criteria: (Motivation AND (Firefighter OR Coast*guard OR Military OR Cadets OR Soldier* OR Mine*workers OR Miners OR Police) AND (Physical Fitness OR Physical Readiness OR Fitness), and yielded 75 hits.

The search on: https://www.san-netz.de/ with the following criteria: (Motivation AND (Physical Fitness OR Physical Readiness OR Fitness) AND (Firefighter*) yielded the following results: PsycARTICLES = 49 hits, PsycBOOKS = 31 hits.

The search on: https://www.san-netz.de/ with the following criteria: (Motivation AND (Physical Fitness OR Physical Readiness OR Fitness) AND (Police*) yielded the following results: PsycARTICLES = 380 hits, PsycBOOKS = 453 hits.

The search on: https://www.san-netz.de/ with the following criteria (Motivation AND (Physical Fitness OR Physical Readiness OR Fitness) AND (Military* OR Soldier OR Cadets) yielded the following results: PsycARTICLES = 602 hits, PsycBOOKS = 712 hits.

A search on November 24th was conducted on Social Science Citation Index with the following criteria: (incentive AND physical performance) and yielded 141 hits. After initial scrutinizing 24 of those were considered.

To include the discussion of gender issues, a search on December 12th was conducted on PubMed with the following criteria: (Motivation AND (Firefighter OR Coast*Guard OR Military OR Cadets OR Soldier* OR Mine*workers OR Miners OR Police) AND (Physical Fitness OR Physical Readiness OR Fitness) AND ((gender differences) OR (sex differences) OR (male vs. female) OR (men vs. women) OR (Gender specific)), which yielded 10 hits. Out of these, our first analyst identified eight as not fitting the current topic. A second analyst considered the two remaining outside topics for further scrutiny. (Note that we were merely looking for research papers on the behavioral aspects of incentives and gender differences; we did omit work on physiological and biological aspects and gender issues from our selective analysis.)

4.2.3 Results

The research papers on promoting health with incentives can be divided in those looking at the effectiveness of single interventions and those which have a broader scope. Several of the single interventions use a pedometer to monitor the physical activity and take steps per day as the measurable performance. Lystrup et al. [20] looked at pedometry as an effective and long-lasting method to increase physical activity. The majority of participants with a pedometer reported increased motivation and physical activity. However, adherence 13 months later was very low. Nevertheless, Garnett et al. [21], looked at a program that promoted physical activity in terms of miles walked and could show that the program got students in their families moving and also improved academic success, student behavior and climate.

As mentioned above, health insurers have a big interest in promoting health conscious behavior and have implemented several bonus systems for their clients. A rather creative malus system was recently implemented by a US public health insurer: it raised out-of-pocket-maximum and the premium for the obese insurant up to 20% unless they started exercising. The latter was monitored by enrollment into a pedometer-based program, which required 5000 steps per day as a minimum [20]. Note that eligibility for participation depended on a specific physical condition, namely being overweight. Forty-three percent (43%) of eligible individuals elected to participate in the program. Among those who did participate, the majority were able to keep the step count above the established goal. Furthermore, almost half of the participants responded in a subsequent survey that the financial incentives were at least partly a motivator for them. Note that financial incentive in this set up could be much larger (under special circumstances amounting to US $2000) than in a classical clinical trial. The authors themselves note this circumstance and state that a clinical trial including incentives of that magnitude would most likely not pass the Institutional Review Board. Aside from these ethical concerns, the authors themselves challenge the generalizability to situations in which incentives are smaller. Of course, if the financial incentives are sufficiently large, they can affect participation and adherence to a program [22], [23]. However, in this context, it should be pointed out that participants did not necessarily receive or save US $2000. The participants could save 20% from their maximum out-of-pocket medical expenses, which could amount to US $2000. If participants did not need services at all, or the cost of their treatments did not reach their out-of-pocket-maximum, savings would be substantially lower.

A clever study, which examined the probability of receiving an incentive in the context of promoting physical activity was conducted by Patel et al. [24]. These researchers also used steps taken as a performance measure, but once the goal was reached, no fixed incentive was awarded. A bet was played. If the bet turned out in the participants' favor, they received its pay-out if and only if the physical exercise goal had been met for the week. The expected value of the bet was US $1.40 per day, but participants could win up to US $350 in a week. This procedure not only allows for a potentially high incentive while still keeping program costs under control, it also operationalizes the uncertainty of outcomes after the investment of physical exercise. That is to say, diligent adherence to a physical exercise program is no guarantee of a long healthy life, it just increases the chances of it. Note that in the studies by Patel et al. [24], [25], [26], participants were also informed if they had won the lottery, but were not eligible to cash in the pay-out because of failing to reach the set exercise goal. This feature adds a component of regret-motivation.

In addition to facing uncertain rather than fixed incentives, participants were in teams of four [27]. The exercise goal was to be met by the team as a collective; there were no minimum requirements for the individual. The teams received feedback as to how their performance compared to the 50th or 75th percentile, thereby leveraging on social comparison. Whether the comparison was expressed to the 50th or 75th percentile did not make a big difference, but the participating teams in the incentive-conditions achieved the exercise goal on significantly more days than the teams in the non-incentive condition. This higher level was maintained in the 13 weeks following the intervention phase. Patel et al. [24], based their design on classical reinforcement theory, which states that the likelihood of a behavior to continue after incentive-removal is greater if the behavior is rewarded intermittently rather than consistently. Also, the study illustrates the efficacy of interdependence by rewarding team, rather than individual, performance.

Adams, Hurley et al. [28], as well as Adams, Sallis et al. [29], examined another aspect in the context of financial incentives as means to motivate physical exercise, namely adaptive (vs. static) goals. Their argument was that smaller more immediate rewards could engage individuals more frequently and thus can more actively shape behavior toward a long-term engagement in physical activity. That is, a very inactive and overweight individual was not required to meet a goal of 10000 steps per day, a goal that they most likely would not meet on many days of the beginning of the intervention. Instead, initially inactive individuals received a very limited challenge, making it much easier to meet the goal. With a limited challenge the likelihood for meeting the goal more often and hence receiving positive feedback from the start of the intervention increased. Adaptive goals can be small and tailored to the individual's prior performance. Step counts over the last nine days were taken and their 60th percentile rounded to the nearest 25 steps became the next day's goal. Results showed that adaptive goals were more effective than static goals with either immediate or delayed rewards. Overall, small, immediate rewards outperformed larger, delayed rewards.

A meta-analysis by Conn et al. [30], is worth mentioning: they looked at physical activity behavior outcomes of 358 reports and coded for 74 intervention characteristics. The authors found that interventions aimed at entire communities were less effective than interventions aimed at individuals. Interventions delivered face-to-face trumped interventions delivered through a technology, i.e., email, phone. Interventions that used behavioral strategies, such as goal setting, cues or rewards, were more effective than interventions that made use of cognitive strategies, e.g., health education, providing health information. Flynn et al. [7], looked at studies examining the impact of culture on health behavior. They looked for one or more of the 24 elements defined by the Health Enhancement Research Organization (HERO) to create a culture of health. Training and learning, policies and procedures, and communication were most frequently evaluated. The authors found that a large number of studies showed a significant relationship between the culture of health elements and the health of employees.

4.2.4 Conclusion

Earlier meta-analyses on the long-term effects of interventions to increase physical activity seem to communicate cautious optimism about the effectiveness of interventions and raise doubts about their cost effectiveness [31], [32]. However, the studies reviewed herein suggest that a careful design of incentives to promote physical activity might be worthwhile. The behavioral elements to be considered in the design of incentives to promote physical activity are:

1) Adaptive intermittent goals, so that requirements can be individually tailored to reach a final goal;
2) Probabilities of incentives, so that higher incentives and uncertainty of pay-outs can be implemented; and
3) Team goals, so that peer pressure and interdependence enhance the benefits of reward.

Many of the reviewed studies were done with overweight and inactive participants. There is reason to believe that the salient features of the design of incentive would work equally well with a different sample. Nevertheless, there should be some caution to generalize the results to a sample of military personnel. Furthermore, as mentioned, many of the reviewed studies used a pedometer, extracting merely steps taken as performance criteria. Additional different parameters could be considered as performance measures in the future.

4.2.5 Workplace Incentives for Physical Activity (Section One Continued)

Studies show that participation in workplace wellness programs increased 40% when incentives are used and over 70% when incentives and penalties are used [33]. A study by Seaverson et al., determined relative contribution to health program participation rates was greatest for incentive value, followed by communications and culture [34]. Most workplace wellness program incentives come in the form of monetary value, Seaverson's study showed that with each $20 of incentive, an increase in participation of 1.6% was shown [34]. RAND Corporation also found that employers offering rewards of more than $100, reported participation rates of 51 percent, compared with 36 percent for those with smaller rewards [33]. Hooker et al., showed that by rewarding subjects with $25 for each month they attended the fitness center at least ten times a month, did increase attendance, but that attendance dropped off after a year [35]. Another study by Pope and Harvey, looked at first-year college age students; they were divided into three groups: no incentive, $5 incentive for each week they attended the fitness centre at least five times for fall semester only, and $5 incentive for each week they attended the fitness centre at least five times for fall and spring semester, and each week they attended the dollar amount increased with a potential to earn $40 a week [36]. After the fall semester, the no incentive group only met the weekly fitness centre attendance goal 13% of the time, while the other incentive groups met it 64% and 62%, respectfully. By the end of the spring semester, the goals were met 3% for both the no incentive and fall only incentive groups, and 39% for the fall – spring incentive group. Another study on financial incentives and weight loss confirmed previous incentive and health behaviour literature suggesting that targeted behaviours do not persist after the incentives are discontinued [37]. When looking at low cost interventions to increase physical activity, Beatty and Katare looked at lottery-based financial incentives and a social norming treatment [38]. The lottery intervention provided a financial incentive to increase physical activity and the social norming intervention attempted to increase physical activity by providing feedback to individuals on their own and their peers' physical exercise activity. They found the larger of two lottery treatments yielded a positive and statistically significant effect on physical activity at modest cost, whereas the social norming treatment had no detectable effect.

There are other initiatives proven to increase physical activity besides monetary incentives. Seaverson's study also showed programs were more efficacious when employing both incentives and communication, versus just one of these [34]. Of the programs using incentives, and regardless of the type of incentive, those that had strong communication strategies had the most participation, increasing it almost 13%. Communication and education are shown to have great impacts on wellness. When looking at university pre-diabetic employees, Abdelsalam and Said showed that employees were unaware they were pre-diabetic and had little knowledge of pre-diabetes risk factors [39]. Subjects were given wellness education classes and monthly support checks for six months. Prior to the education 21% knew diabetes risk factors, after the classes 77% knew, and while only 6% had a normal health profile before the classes, healthy profiles increased to 63% after classes and monthly support follow ups. Technology support may also be an avenue of increasing exercise. The results of one study [40] suggested that text message reminders of member's implementation intentions and SMS reminders increased exercise frequency significantly, while another study conducted a combination of podcasting, mobile support communication, mobile public announcements and diet to assist people in weight loss, also showing positive results [41]. Finally, communication and leadership support were contributing factors to "buy-in" in a study of wellness program uptake in Fire Service Organizations [42]. The study determined that programs that had the best "buy-in" were those that had leadership support (both general administration and leading by example categories). The next two sections address how military services can align incentives across two major types of physical fitness tests and section four addresses incentives that military services could consider and apply to their current and future physical fitness tests.

4.3 SEX-FAIR PHYSICAL FITNESS TESTS AND STANDARDS

4.3.1 Delineation of Sex-Fair Physical Tests and Standards to Which Military Services May Apply Incentives

Promoting health and well-being is addressed above, but how best may military services align incentives to physical employment standards? Prior to aligning incentives to physical fitness tests and standards or physical fitness training a military service should communicate to military members the criterion basis for the standards. This can aid in program face validity and acceptance of the incentives, i.e., if members perceive an age, sex, or other bias in the standards then the incentives will likely be deemed biased as well. Consensus recommendations from exercise science subject matter experts at a US Air Force workshop on exercise compliance include the need to develop and communicate to leaders and members health-related and occupation-related fitness test standards and exercise adherence initiatives for both [43]. The USAF delineates its physical fitness test standards across two science-based “tiers”; Tier 1 health and general fitness, and Tier 2 Occupationally-Specific, Operationally-Relevant (OSOR) [44] [45]. USAF Tier 1 standards are sex dependent, for the *same health outcome* men and women are at different levels of fitness; the corollary, for the same level of fitness men and women have different health outcomes, different health risk levels. In this fitness test a male must run faster than a female to achieve the same health risk, yet a female must possess a lower abdominal circumference measure than a male to achieve the same health risk. When considering incentives for this Tier 1 health-related test as well as a Tier 2 OSOR test the key is to align the incentive with the criterion outcome. The USAF accomplishes this by linking an incentive (altered test frequency) to the Tier 1 component point score as the scores are directly derived from the level of health risk, the criterion outcome [45], [46]. For military fitness tests with sex-fair normative standards, without a criterion basis, the incentive alignment may work ***if*** members perceive the standards as unbiased. This is not always the case with these “gender-normed” test standards. This concern is not a factor with criterion-based tests, e.g., Tier 1 above, or when considering sex independent or sex-free standards as one finds in the following OSOR tests.

4.4 SEX-FREE PHYSICAL FITNESS TESTS AND STANDARDS

4.4.1 Delineation of Sex-Neutral Physical Tests and Standards to Which Military Services May Apply Incentives

This section opens with a description of incentive strategy in the Canadian Defence Forces fitness tests which are based on health-related standards (sex-fair) and operational-related military tasks (sex-free) and concludes with a description of linking incentives to the US Air Force’s Tier 2 OSOR criterion physical fitness tests and standards (sex-free).

4.4.2 Incentive Strategy of your Organisation

Fitness for Operational Requirements of Canadian Armed Forces Employment (FORCE) is a field expedient fitness test designed to predict the physical requirements of completing six common military tasks. The standards are age-free and sex-free. Although referred to as a “Fitness Test,” it was never intended to be a general test of health-related fitness [47]. In general terms, physical fitness is defined by the components of fitness demonstrated in Figure 3-3 (Chapter 3), aerobic capacity, anaerobic capacity, muscular strength, muscular endurance, muscular power, flexibility, balance, speed, agility and coordination.

Although it may seem peculiar that a military fitness test would not contain a strong aerobic component, the six military tasks on which the FORCE evaluation is based tend to be low intensity, long duration, and manual material handling tasks, which do not elicit a high aerobic response [48], similarly a recent PES review of 57 employment categories within the Australian Army identified 583 physically demanding tasks, of which 458 (~79%) were classified as manual handling tasks [49]. Although, the FORCE evaluation includes elements of muscular strength, endurance, and flexibility it does little to reflect body composition and cardiovascular endurance. Given that attaining this minimal physical employment standard may not represent a challenge to some personnel, a fitness incentive program was requested by the chain of command to recognize and reward physical and occupational fitness over and above the minimal standard, which is measured at an oxygen consumption of 26.3 ml $O_2 \cdot kg^{-1} \cdot min^{-1}$, and is therefore not sufficient to indicate good aerobic fitness, or protect from morbidity and mortality.

In the past decade various employers have adopted a "Health-Based" fitness approach and applied a cardiovascular fitness standard to reflect the decrease in risk of morbidity, mortality and injury associated with increased fitness [50], [51]. The US Army National Guard employs a 2 mile run and categorises personnel on the basis of a 10 year Coronary heart disease risk [52]. This was a result of identifying that their deployed members were experiencing CVD symptoms despite being screened as low or moderate risk on the Framingham profile (469 cardiac referrals of deployed soldiers with a mean age of 39 years) [52]. The US Air Force developed a scoring system on their fitness test creating a health profile which is determined based on cardiovascular risk, and risk of injury, assessed through a 1.5 mile run, push-ups, sit-ups and abdominal circumference [45]. Similar emergency services such as The Royal National Lifeboat Institution in the United Kingdom [53], and Firefighters in the USA and Canada [54] have a cardiovascular fitness standard with the intention of improving health, while not always directly linked to the metabolic demands of the job, but more the categorisation of "fit" and "unfit."

In order to encourage health-related benefits, the CAF research team employed the foundation of research linking low Cardiorespiratory Fitness (CRF) to increased mortality and morbidity [55], [56], [57], [58]. This relationship has been shown to be independent of age, ethnicity, smoking status, alcohol intake, and health conditions [59]. Several authors have therefore attempted to quantify and categorise healthy and unhealthy VO_{2max} cut-scores for males and females of different ages [55]. Although these vary slightly in the literature, an extensive meta-analysis by Kodoma et al. [56], recently proposed the following cut-scores for unhealthy CRF, based on relative risks of all-cause mortality and morbidity [56] (Table 4-1).

Table 4-1: High and Low Risk Cut-Offs for 40 Years – 60 Years of Age Based on Ref. [56] and Extrapolated for Those Under 40 Years [48]. VO_{2max} ($mlO_2 \cdot kg^{-1} \cdot min^{-1}$).

	Male		Female	
Risk Cut-off				
Age	High	Low	High	Low
15 – 20	38.1	48.6	31.5	41.7
20 – 25	36.4	46.9	29.4	39.9
25 – 30	34.6	45.1	27.6	38.1
30 – 35	32.9	43.4	25.9	36.4
35 – 40	31.1	41.6	24.1	34.6
40 – 45	29.4	39.9	22.4	32.9
45 – 50	27.6	38.1	20.6	31.1
50 – 55	25.9	36.4	18.9	29.4
55 – 60	24.1	34.6	17.1	27.6

From these relationships a VO_{2max} below 9, 8, or seven Metabolic Equivalents (METS) for males aged 40, 50, and 60 years, respectively, was considered unhealthy. For females, a VO_{2max} below 7, 6, or 5 METS was considered unhealthy aged 40, 50, and 60 years, respectively.

Body composition, as measured by Abdominal Circumference (AC), is also linked to long-term health outcomes [60], [61], [57], [62], [63]. Leading regulatory bodies recommend annual monitoring of AC by medical personnel [64]. It has been shown that AC has a stronger relationship to heart disease risk when compared to percent body fat, body mass index, and waist-to-hip ratio [55]. Because of a failed policy administration in the early 1990's with the CAF, using body mass index led to questionable dismissals of military personnel from service. For that reason, many in the CAF leadership and medical community believed that introducing a measure of AC as part of the annual FORCE evaluation would trigger apprehension in the general CAF population. This fear was alleviated by making it clear to participants that the AC measurement in the fitness profile is not linked to any employment condition or policy [47].

After the scoring tables by age and/or sex, based on the risk ratios of chronic disease for both VO_2 and AC, were developed (Table 4-1), these were weighted at 75% (VO_2) and 25% (AC) of the health-related physical fitness score to account for the buffering influence that high aerobic fitness has on the detrimental health effects of obesity [52].

To predict relative maximum aerobic capacity (VO_{2max} ml $O_2 \cdot kg^{-1} \cdot min^{-1}$) 195 male and female military members, 17 – 59 years of age, with a wide range of fitness and anthropometric measurements, were recruited from the National Capital Region, 14 garrison Petawawa, and the Canadian Forces Leadership and Recruit School (CFLRS). On two separate testing days all participants performed:

1) A maximal Graded Exercise Treadmill test (GXT); and
2) A maximum effort FORCE evaluation, defined as their best effort [65].

Linear regressions were performed to identify the criterion of measured relative VO_{2max}, using Stepwise and Enter methods. The predictive ability of the models all demonstrated coefficients of determination (R^2) above 0.72, and so are considered reasonable predictions for estimating CRF [66]. These regressions yielded errors (%SEM) ranging from 8.4% to 9.1%, which fall within the 10% error range accepted for predictive tests [67]. This ability to predict VO_{2max} from the FORCE evaluation provided the means to determine a score of occupational fitness (the Y axis, Figure 4-1) and health-related physical fitness [75% (VO_2) and 25% (AC)] (the X axis, Figure 4-1).

Figure 4-1: CAF Fitness Profile; Operational Fitness (Vertical Axis) is Based on Performance of the Four FORCE Tasks, Health-Related Fitness (Horizontal Axis) is Calculated Based on Estimated Cardiorespiratory Fitness and WC [48].

The resulting incentive program structure is now based on sex and eight age categories. The results on the four elements of the FORCE evaluation were converted to a point scale from which normative scores were derived, where the median score corresponds to the bronze level, and silver, gold, and platinum correspond to a score which is one, two and three standard deviations above this median, respectively.

A suite of rewards including merit board points toward promotions and recognition on the uniform and material rewards was developed. A separate group rewards program was also presented, to recognize achievements in fitness at the unit level. Pilot testing of this entire program was performed with 624 participants to assess participants' reactions to the enhanced test, and also to verify logistical aspects of the electronic data capture.

Although the motivating effect of the suite of rewards attached to performance on the fitness profile cannot be assessed for several years, it certainly represents the integration of a solid theoretical grounding with items perceived to be valuable by CAF personnel [47]. The resulting rewards are material in nature, and not linked to promotion, pay or annual leave.

4.4.3 Link Incentives to OSOR Standards

Physical employment test standards with performance outcome criteria found in OSOR tests such as the Canadian FORCE test above or the USAF's Tier 2 test are sex independent [47], [48], [44], [45]. Therefore, incentives linked to the criterion outcome of occupational task achievement are sex-free, men and women must achieve the same fitness level that predicts or is linked directly to a fixed occupational performance outcome. These role-related standards and associated incentives have high content and face validity, they are usually readily understood and accepted by military personnel of both sexes. OSOR test standards may be pass fail or as the USAF employs and recommends a scoring scale, with criterion values above the minimum, which acts as an objective basis to assign ever increasing incentives to motivate members to achieve and maintain fitness levels above the minimum requirements [45], [68], [69], [70]. As a USAF Airman improves in any one or more of the ten fitness components depicted in the Tier 2 OSOR scoring chart at Table 4-2, that Airman has an ever increasing probability of successfully accomplishing the operationally-relevant physical tasks necessary for combat success [71].

Table 4-2: USAF Operational Level Tier 2 Physical Fitness Test Standards for Tactical Air Control Party Airmen.

ALO-TACP Points	Grip Strength PSI	Med Ball Toss ft	Two Cone Drill secs	Trap Bar DL 5RM lbs	Pull up reps	Lunges Wtd reps	Ext Cross Knee Crunch reps	Farmer's Carry 4 x 25 yd secs	Row 1000 m min:secs	Run 1.5 mile min:secs
10	198	50.5	8.6	417	32	199	107	21.2	3:15	7:51
9	166	47.5	8.9	387	28	161	94	22.5	3:22	8:03
8	153	44.5	9.3	355	24	148	79	23.5	3:30	8:50
7	144	42.0	9.5	334	22	123	69	24.8	3:36	9:21
6	137	40.5	9.8	316	20	104	62	25.9	3:40	9:47
5	130	39.0	10.0	300	18	88	57	26.8	3:44	10:10
4	124	37.5	10.2	284	16	73	52	27.8	3:48	10:33
3	117	35.5	10.4	267	13	58	46	28.8	3:52	10:59
2	108	33.5	10.7	245	11	43	39	30.1	3:57	11:31
1	95	30.0	11.1	213	7	24	31	32.0	4:05	12:17
ALO-TACP - Component Minimums indicated in blue						Composite score requirement ≥ 46 of 100				

The ten by ten score chart that provides the platform for assigning incentives also provides additional incentive to the member by:

1) Indicating current fitness level across all test components and in turn multiple physical fitness components;
2) Provides training staff the feedback to guide or redirect individual physical training; and
3) Provides leadership with the physical strengths and weaknesses of this specific team member (Table 4-3).

Applying incentives to a Tier 2 OSOR test should be straightforward as these test standards are independent of sex, age, rank, and ethnicity (linkages to employment law varies for some nations). It is agreed across many international military organisations that if a military member fails to meet minimal physical fitness test requirements, even after a period of rehabilitation and retesting, they should be removed from the career field or discharged; however, some nations may not legally expect a member to exceed the minimum standard.

Table 4-3: Individual Data from an Individual USAF Battlefield Airmen on a USAF Operational Level Tier 2 Physical Fitness Test.

PJ/ CRO	Grip Strength Max	Med Ball Toss Sum	Three Cone Drill	Trap Bar DL	Pull up	Lunges Wtd	Ext Cross Knee Crunch	Farmer's Carry	Row Erg 1000m	Run 1.5 Mile
Points	PSI	ft	secs	lbs	reps	reps	reps	secs	mins:secs	mins:secs
10	195	50.5	7.4	467	35	185	168	13.0	3:07	7:40
9	161	46.5	7.7	403	30	152	149	14.5	3:18	8:15
8	149	43.5	8.1	370	26	120	130	16.0	3:28	9:03
7	140	41.0	8.4	348	23	100	111	17.7	3:36	9:35
6	134	39.5	8.7	330	21	86	92	19.2	3:41	10:01
5	128	37.5	8.9	314	18	74	73	20.5	3:46	10:26
4	121	36.0	9.1	298	16	62	55	21.8	3:52	10:49
3	115	34.0	9.4	280	14	51	40	23.3	3:58	11:16
2	106	32.0	9.7	258	11	39	28	25.0	4:05	11:47
1	94	28.5	10.1	226	7	24	16	27.7	4:16	12:35

Component Minimums indicated in blue

Composite score requirement ≥ 52 of 100

Also, as described in the chapter on biological differences, the average female military member would have to undergo more extensive physical training than the average male member to achieve the same absolute Tier 2 OSOR physical standard as depicted in Figure 4-2. Females may require a more focused, consistent, periodized training program than males for cardiorespiratory fitness, body composition and muscular fitness, e.g., females generally need to achieve greater improvements in cardiorespiratory fitness and muscular fitness to reach the same absolute load carriage capability [72].

To summarize Sections 4.2 and 4.3, fairness can be achieved by aligning incentives to the fitness test score (or change in score) when the score is aligned to a criterion standard. In a health-based test the incentive should be aligned to the test score linked to the criterion health risk, in an occupationally based test the incentive should be aligned to the criterion operational performance. For example, a USAF female Airman running 14:05 minutes:seconds on the 2414 meter run test has an estimated VO_2 max of 37 ml $O_2 \cdot kg^{-1} \cdot min^{-1}$ achieving a low health risk, the same level of health risk a male achieves by running the 2414 meter test in

11:30 minutes:seconds, an estimated VO_2 max of 45 ml $O_2 \cdot kg^{-1} \cdot min^{-1}$. However, the female Airman would have to increase her cardiorespiratory fitness to run 11:30 minutes:seconds, an estimated VO_2 max of 45 ml $O_2 \cdot kg^{-1} \cdot min^{-1}$, to reach the same operational performance required to pass the aerobic component of the USAF Tier 2 Tactical Air Control Party OSOR Physical Fitness Test. Incentives should be aligned accordingly in Tier 1 and Tier 2 tests.

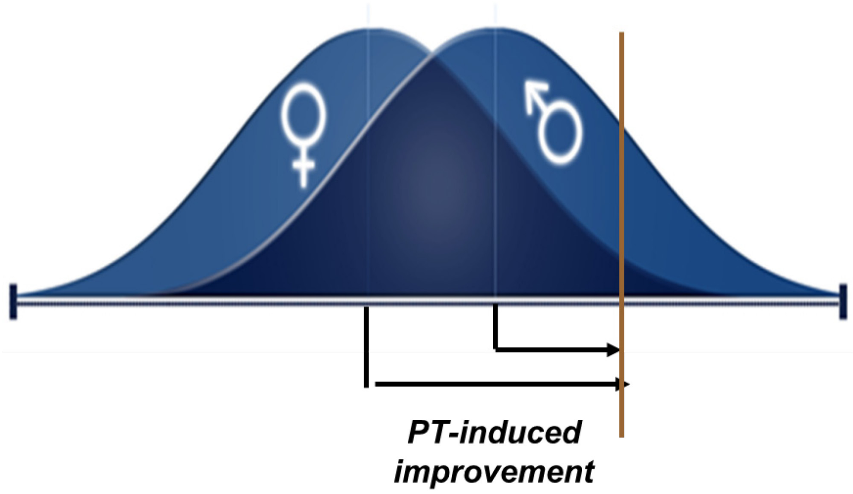


Figure 4-2: Sex-Specific Differences Lead to Different Physical Training Requirement.

4.5 MILITARY-SPECIFIC LIMITATIONS AND OPPORTUNITIES TO INCENTIVIZE PERSONNEL

Finally, section four addresses incentives that military services could consider and apply to motivate their military members to meet and then exceed minimum physical fitness test standards and increase quantity and quality of physical training.

4.5.1 Incentives in Military Physical Fitness Training and Testing

Incentive use in military physical fitness testing and training is generally not extensive, for example the USAF Physical Fitness Program has penalties built into the regulation by having members test more frequently if they score below an excellent, and derogatory marks on performance reports for repetitive failures, but there isn't much in the way of incentives built into the program. Therefore, we address a myriad options that military services may consider for use across ten different categories in this section. Employing a balance of "carrot and stick" approach as addressed in section one above (bonus and malus) is supported [33] and encouraged [70], [73] in the literature.

4.5.1.1 Monetary

Monetary incentives are often used in civilian settings as addressed above and are the second ranked choice of USAF Airmen in an ongoing 2019 physical fitness survey [74], however, military services may not be permitted to apply monetary incentives to service members, and other legal limitations may come in to play as well. Alternatives might be the use of morale, welfare, recreation funding for physical activity incentives, e.g., awards listed in Table 4-4. Also, services that already award incentive pays for certain career fields for special duties such as parachute free fall or underwater diving, could consider monthly incentive pay for members that score a specified high level on their fitness test and maintain that high level of fitness over protracted time periods.

Table 4-4: Incentives for Physical Fitness Testing and Training.

Incentive	Notes
MONEY	
Specialty pay	
Bonus	
Tax incentives	
Life insurance premium reduction	Running and heart protection; run an 8 minute mile? [75]
Health care benefits	Lower premiums, free vaccinations
TIME	
Time off / passes	
Duty time for training	
Early departure / Late show	
FITNESS TESTING	
Test high score	
Test most improved	
Test score maintained	Over time (years)
Test frequency, altered	Inverse to test score
Random testing	Control variables to minimize fear and resource impact
Test anxiety	Prior to official test offer a “free” test: pass, it counts officially / fail, no consequence (still retake officially scheduled test)
Scoring scale above minimum P/F	More than the “minimum”
FITNESS TRAINING	
Participation in physical training	Reward volume and intensity of training
Expanded fitness centre access hours	Wider time window for training during day
Single sex training venue / Fear of the gym	
Fitness pairing with workout partners	
AWARDS	
Patches	PT gear
Badges	PT gear

Incentive	Notes
Ribbons	Military uniform
Medals	Military uniform
Promotion points	
Numerical score on performance reports	Current test score and average of last 3 years
Unit awards	Measures of multiple members exercising
Recognition, amount of exercise	Letters of recognition, point clubs – running, cycling, rowing distance covered, mass lifted
Rank in unit or on team	
Physical activity gift	Entry fee to fun run; promote physical activity
Base of preference	When reassigned
Choice of military courses	
Incentive rides	Unique to military, e.g., aircraft ride
Parking spot	
Free photo shoot	
Casual dress	
Office competition / raffles	
ENVIRONMENT	
Career field requirements	Scoring above minimal requirements
Work place environment	Stairs, standing desks
Peer pressure	
Team and individual; unit performance	Competitions, raffles; unit status within larger organisation
Fear of Missing Out (FOMO)	Positive opportunities and group influence
Absenteeism	Inverse relationship to rewards
Multi-factor member feedback	Fitness and health assessment combined
LEADERSHIP	
Mentoring and support	
Example	
EDUCATION	
Education, rationale	Basis of fitness test and program

Incentive	Notes
Education, physical training	
Fitness course / certification offerings	
Personal trainer, access to	
TECHNOLOGY	
Funding for fitness wearables	GPS watches, other physical activity devices
Health-related apps	
Computer notifications / warnings	Prevent protracted sitting, exercise reminders
BEHAVIORAL STRATEGIES	
Journaling, charting	
Wellness check-ups	
Social media	
Social comparison	
Social rewards	
Goal setting	

4.5.1.2 Time

Since monetary incentives are not an option for most military leaders, another desirable incentive is time. Time off passes are favoured over ribbons by USAF Airmen [70]. Leaders are able to give their members time, a day off for scoring a specific level on the fitness test or time off during the work day to execute physical fitness training. Some military service fitness regulations stipulate that members be allowed a certain amount of time during the work week for physical training, yet a vast number of members do not take fitness time during the duty day. A possible reason for this is the stigma of being observed as slacking on primary duties if one takes duty time to exercise. Commanders can alleviate this issue by setting an example by using the allotted time during the day to show the importance of fitness training to their subordinates. In a study of work place fitness programs among fire fighters, Fire Chiefs exercising with employees during duty day was a leading employee motivator [42]. Finally, mandating physical fitness during duty time versus offering duty time as an option remains equivocal amidst military members [73].

4.5.1.3 Fitness Testing

Military services have another time variable in their control and that is physical fitness test frequency. Incentive rewards and penalties may be applied by inversely linking test frequency to fitness test scores with the goal of altering physical fitness training behavior, or exercise compliance, towards greater consistency [73]. In an ongoing 2019 physical fitness survey the USAF found that 67% of Airmen (n = 553) of all ranks agreed and only 17% disagreed the USAF should incentivize Airmen by lowering test frequency for higher fitness test scores [74]. A lower test frequency will also reward some who experience test anxiety as pre-fitness test anxiety demotivates many Airmen, even those that have no issues passing the test. A lower test frequency may help quell this fear as well as a “free” test, i.e., offering an opportunity approximately 30 days to 45 days prior to an Airman’s official fitness test to take a test, if pass, then it counts officially, if fail, no consequence (but, still retake officially scheduled test). In the same survey 76% of Airmen agreed or strongly agreed and only 17% disagree or strongly disagreed with this concept. The USAF is seeking a balance between incentives, accountability, and test resources. Therefore, the USAF Exercise Science Unit has undergone initial work to apply altered test frequencies to motivate Airmen and save test resources as nearly 60% of Airmen score > 90 points out of a maximum of 100, extending their next test to 15 months or 18 months vice the current 12 months would save over 160,000 tests per year. Additionally, to enhance accountability the USAF is considering controlled (limited number and opportunity) random testing, which had a strong consensus support in 1999 [73] and 2001 [70] amongst exercise subject matter experts, and with Airmen in 2019, although to a somewhat lesser degree [74].

4.5.1.4 Fitness Training

Applying incentives to fitness testing is most common in civilian and military settings, however, motivating military members via incentives to adhere to consistent physical fitness training as a lifestyle across their career is of greater import, and efficient training that takes the least amount of time per exercise session will likely improve exercise compliance [69]. Other training incentives as fitness pairing with workout partners, group exercise sessions, and training participation rewards (Table 4-4) [71]. Additionally, to incentivize Airmen who fear gym attendance for a variety of reasons, military services could offer single sex venues or time periods at fitness centres in addition to fitness educational opportunities (see Section 4.5.1.8 below).

4.5.1.5 Awards

Numerous award incentives are possible, some are listed in Table 4-4. A useful incentive would be activity oriented rewards or passes, e.g., a trip to an outdoor recreational activity to promote members' physical activity vice common passivity watching others (attending sporting event as a fan) [70].

A well discussed incentive is linking physical fitness scores to promotion. Services could reward fitness test results – high scores, most improved, maintained high over multiple years – with promotion points and include the results on performance reports [74]. Military services also have the unique opportunity to reward members with military-specific incentive rides to members earning positive test or training results, e.g., incentive aircraft rides, shooting ranges, defensive driving courses, and outdoor recreational events.

4.5.1.6 Environment

Military services should also enhance the work environmental to support and incentivize physical training [70] with safe and accessible stairs, military fitness centres, exercise classes, shower/changing facilities, walk-run trails, playing fields [43]. With the strong linkage between fitness scores and health care costs, absenteeism, and military readiness private score presentation to individuals and open presentation of aggregate unit scores versus overall organisational scores showing fitness and other readiness factors such as immunizations, sick days used, may also be considered [70].

4.5.1.7 Leadership

As above, in the category on time, leadership support and leadership by example can be incentives. The USAF does an adequate job in the general administrative support of fitness (time off during the workday, facilities), there remains large deviations in the lead by example category. Most commanders and leaders feel pressure to do more with less, their days fill quickly with taskers and meetings, leaving little time for subordinates to witness leadership taking time to exercise on duty. This in turn, puts an unconscious bias on subordinates to stay and work on taskers and not use the duty time for fitness offered.

4.5.1.8 Education and Communication

Military services should clearly communicate the rationale behind fitness tests and programs, this a simple incentive for members' attitudes towards testing and training, for example the USAF and the CAF have science-based health risk categories indicated in green (low risk), amber (moderate risk), and red (high risk) across the fitness test score range, however the USAF does not print score charts in color, inhibiting the communication of this simple rationale. Educational resources to market exercise benefits in pipeline training and throughout the military career [70], [73] and education, supervision and encouragement early in a physical training program are important to exercise adherence, along with maintaining solid communication and follow-up [43]. Incentives for test administrators, fitness trainers [70] and physicians clearing members to exercise are also important. Despite demands placed on them regarding physical fitness (clearing members to test and train) medical providers may have limited knowledge of physical fitness and health relationship [76]. Educational "road shows" on physical fitness can go beyond electronic means to effectively incentivize members [70], but electronic interventions such as emailing physical fitness and nutrition interventions to military personnel are also effective [77].

4.5.1.9 Technology

Interventions as above [78] make use of office computers that may also be used to notify members when screen time becomes protracted and a notification can incentivize members to stand or depart for the fitness centre. Also, the expanding market for fitness wearables and applications offer opportunities for incentive rewards.

4.5.1.10 Behavioural Strategies

Finally, journaling one's fitness actions and progress may be an incentive as well as social comparison of fitness testing and training via social media. Finally, recent USAF survey data (n = 160) on incentives are at Table 4-5 [74].

Table 4-5: USAF Physical Fitness Survey.

Which incentives would you like to see as motivational tools for positive physical fitness testing and training behaviors *(select all applicable)?*			
	Answers – Rank Order	**%Voters**	**%Votes**
1	Time off Pass	75%	30%
2	Monetary Reward	47%	19%
3	Ribbon (Service Dress)	31%	13%
4	Qualify for Professional Fitness Course	30%	12%
5	PF Patch/Badge	24%	10%
6	Fitness Leader Eligibility	22%	9%
7	Other	18%	7%
	Members could select more than one of the choices		
	%Voters = percent of all voters selecting this choice		
	%Votes = percent of total number of votes		

4.5.2 New Audience / Know Your Audience

The majority of military members is now made up of a "different generation", millennials 22 years to 38 years of age. Their motivations are different than the generation before them. Millennials aren't driven by money or the company's bottom line. They are motivated by the company's mission, it needs to invigorate and inspire, and they are drawn to inspirational CEOs, ones that embody the company mission and values, ones that speak with authenticity and conviction and engage employees at all levels. How does this translate to the military? The same way the Fire Fighter study [42] found, to motivate employees leaders need to lead by example when it comes to the fitness aspect of duty and not just the "Do as I say" or "Do it because it's your job" mentality [78]. Younger employees are also more competitive, a recent study [79] found that nine out of ten millennials today consider themselves competitive with 80% directly stating they are better than a co-worker. Hence unit fitness competitions should work as incentives and rewards should be less material and more experiential and social. A recent study found that 72 percent of millennials prefer to spend more money in the next year on experiences than on material items, pointing to a move away from materialism and a growing appetite for real-life experiences [79]. Another aspect to rewarding experience is creating a social media buzz about it. Nearly seven in ten millennials experience FOMO (Fear Of Missing Out). In a world where life experiences are broadcasted across social media, the fear of missing out drives millennials to show up, share and engage; 79% of millennials feel that going to live events with family and friends helps deepen their relationships and 69% of millennials believe attending events makes them feel more connected to other people, the community, and the world [79].

4.6 SUMMARY

In summary the extensive literature on physical activity incentives in the workplace provides a background for incentive methods and successes which military services could consider. If applying incentives to physical fitness tests military services should link incentives to sex-fair or sex-free criterion standards that are well understood and received by personnel. Finally, several incentive methodologies are available to meet military-specific limitations and opportunities with the overall goal to motivate military personnel to meet or exceed physical fitness standards and, most importantly, conduct consistent physical fitness training.

4.7 REFERENCES

[1] Herbermann, J.D. (2018). Ein Bauer kaempft fuer die Hoerner der Kuehe [A farmer's battle for cows' horns]. Retrieved from https://www.tagesspiegel.de/weltspiegel/schweiz-ein-bauer-kaempft-fuer-die-hoerner-der-kuehe/23618176.html.

[2] Ramnath, S. (2013). Taxpayers' Responses to Tax-Based Incentives for Retirement Savings: Evidence from the Saver's Credit Notch. Journal of Public Economics, 101, 77-93.

[3] Gruber, J and Wise, D. (2000). Social Security Programs and Retirement Around the World. Research in Labor Economics, 18, 1-40.

[4] Baicker, K., Cutler, D., and Song, Z. (2010). Workplace Wellness Programs Can Generate Savings. Health Affairs, 29(2), 304-311. https://doi.org/10.1377/hlthaff.2009.0626.

[5] Clayton, G., Rae, M., Long, M., Damico, A., and Whitmore, H. (2018). Health Benefits in 2018: Modest Growth in Premiums, Higher Worker Contributions at Firms with More Low-Wage Workers. Health Affairs, 37(11), 1892-1900. https://doi.org/10.1377/hlthaff.2018.1001.

[6] Krankenkasse.de (2018). Gesundheitsförderung – Bonusprogramme: Von Gesundheitskursen über Geldprämien bis Sachleistungen [Promoting Health – Bonus programs: On health education, monetary incentives and other benefits] Retrieved from https://www.krankenkassen.de/gesetzliche-krankenkassen/leistungen-gesetzliche-krankenkassen/praevention-vorsorge-krankenkassen/.

[7] Flynn, J. P., Gascon, G., Doyle, S., Matson-Koffman, D.M., Saringer, C., Grossmeier, J., Tivnan, V., And Terry, P. (2018). Supporting a Culture of Health in the Workplace: A Review of Evidence-Based Elements. The American Journal of Health Promotion, 32(8), 1755-1788. https://doi.org/10.1177/0890117118761887.

[8] Leyk, D., Rohde, U., Hartmann, N. D., Preuss, P. A., Sievert, A., And Witzki, A. (2014). Results of a Workplace Health Campaign. What Can be Achieved? Deutsches Ärzteblatt International, 111(18), 320-327. https://doi.org/10.3238/arztebl.2014.0320.

[9] Fries, J. F., Harrington, H., Edwards, R., Kent, L.A., and Richardson, N. (1994). Randomized Controlled Trial of Cost Reductions from a Health Education Program: The California Public Employees' Retirement System (PERS) Study. The American Journal of Health Promotion, 8(3), 216-223. https://doi.org/10.4278/0890-1171-8.3.216.

[10] Leigh, J.P., Richardson, N., Beck, R., Kerr, C., Harrington, H., Parcell, C.L., and Fries, J.F. (1992). Randomized Controlled Study of a Retiree Health Promotion Program. The Bank of American Study. Archives of Internal Medicine, 152(6), 1201-1206.

[11] Ozminkowski, R.J., Dunn, R.L., Goetzel, R.Z., Cantor, R.I., MurnanE, J., and Harrison, M. (1999). A Return on Investment Evaluation of the Citibank, N.A., Health Management Program. The American Journal of Health Promotion, 14(1), 31-43. https://doi.org/10.4278/0890-1171-14.1.31.

[12] Griffin, S.C., Regan, T. L., Harber, P., Lutz, E.A., Hu, C., Peate, W.F., and Burgess, J.L. (2016). Evaluation of a Fitness Intervention for New Firefighters: Injury Reduction and Economic Benefits. Injury Prevention, 22(3), 181-188. https://doi.org/10.1136/injuryprev-2015-041785.

[13] Poplin, G.S., Roe, D.J., Peate, W., Harris, R.B., And Burgess, J.L. (2014). The Association of Aerobic Fitness with Injuries in the Fire Service. The American Journal of Epidemiology, 179(2), 149-155. https://doi.org/10.1093/aje/kwt213.

[14] Poston, W.S.C., Haddock, C.K., Jahnke, S.A., Jitnarin, N., and Day, R.S. (2013). An Examination of the Benefits of Health Promotion Programs for the National Fire Service. BMC Public Health, 13, 805. https://doi.org/10.1186/1471-2458-13-805.

[15] Mayer, J.M., Nuzzo, J.L., And Dagenais, S. (2013). Use of Participant Focus Groups to Identify Barriers and Facilitators to Worksite Exercise Therapy Adherence in Randomized Controlled Trials Involving Firefighters. Patient Preference and Adherence, 7(1), 207-215. https://doi.org/10.2147/PPA.S42507.

[16] Leyk, D., Rohde, U., Gorges, W., Ridder, D., Wunderlich, M., Dinklage, C., Sievert, A., Rüther, T., and Essfeld, D. (2005). Adipositas und Bewegungsmangel in Deutschland: Erste Fakten aus der "Physical-Fitness-Test" Studie. Wehrmedizinische Monatsschrift, 49(1), 11-15.

[17] Leyk, D., Rohde, U., Gorges, W., Ridder, D., Wunderlich, M., Dinklage, C., Sievert, A., RütheR, T., and Essfeld, D. (2006). Physical Performance, Body Weight and BMI of Young Adults in Germany 2000 – 2004. Results of the Physical-Fitness-Test Study. The International Journal of Sports Medicine, 27(8), 642-647. https://doi.org/10.1055/s-2005-872907.

[18] Niebuhr, D.W., Scott, C.T., Powers, T.E., Li, Y., Han, W., Millikan, A.M., and Krauss, M.R. (2008). Assessment of Recruit Motivation and Strength Study. Preaccession Physical Fitness Assessment Predicts Early Attrition. Military Medicine, 173(6), 555-562.

[19] Santtila, M., Kyröläinen, H., Vasankari, T., Tiainen, S., Palvalin, K., Häkkinen, A., and Häkkinen, K. (2006). Physical Fitness Profiles in Young Finnish men During the years 1975 – 2004. Medicine and Science in Sports and Exercise, 38(11), 1990-1994. https://doi.org/10.1249/01.mss.0000232023.28984.78.

[20] Zulman, D.M., Damschroder, L.J., Smith, R.G., Resnick, P.J., Sen, A., Krupka, E.L., and Richardson, C.R. (2013). Implementation and Evaluation of an Incentivized Internet-mediated Walking Program for Obese Adults. Translational Behavioral Medicine, 3(4), 357-369. https://doi.org/10.1007/s13142-013-0211-6.

[21] Garnett, B.R., Becker, K., Vierling, D., Gleason, C., Dicenzo, D., and Mongeon, L. (2016). A Mixed-Methods Evaluation of the Move it Move it! Before-School Incentive-Based Physical Activity Programme. Health Education Journal, 76(1), 89-101. https://doi.org/10.1177/0017896916652432.

[22] Charness, G. and Gneezy, U. (2009). Incentives to Exercise. Econometrica, 77, 909-931.

[23] Gneezy, U. and Rustichini, A. (2000). Pay Enough or Don't Pay at All. Quarterly Journal of Economics, 115, 791-810.

[24] Patel, M.S., Volpp, K.G., Rosin, R., Bellamy, S.L., Small, D.S., Fletcher, M.A., Osman-Koss, R., Brady, J.L., Haff, N., Lee, S.M., Wesby, L., Hoffer, K., Shuttleworth, D., Taylor, D. H., Hilbert, V., Zhu, J., Yang, L., Wang, X., and Asch, D.A. (2016). A Randomized Trial of Social Comparison Feedback and Financial Incentives to Increase Physical Activity. The American Journal of Health Promotion, 30(6), 416-424. https://doi.org/10.1177/0890117116658195.

[25] Patel, M.S., Asch, D.A., Rosin, R., Small, D.S., Bellamy, S.L., Eberbach, K., Walters, K.J., Haff, N., Lee, S.M., Wesby, L., Hoffer, K., Shuttleworth, D., Taylor, D. H., Hilbert, V., Zhu, J., Yang, L., Wang, X., and Volpp, K. G. (2016). Individual Versus Team-based Financial Incentives to Increase Physical Activity. A Randomized, Controlled Trial. The Journal of General Internal Medicine, 31(7), 746-754. https://doi.org/10.1007/s11606-016-3627-0.

[26] Patel, M.S., Asch, D. A., Rosin, R., Small, D.S., Bellamy, S.L., Heuer, J., Sproat, S., Hyson, C., Haff, N., Lee, S. M., Wesby, L., Hoffer, K., Shuttleworth, D., Taylor, D.H., Hilbert, V., Zhu, J., Yang, L., Wang, X., and Volpp, K.G. (2016). Framing Financial Incentives to Increase Physical Activity Among Overweight and Obese Adults: A Randomized, Controlled Trial. Annals of Internal Medicine, 164(6), 385-394. https://doi.org/10.7326/M15-1635.

[27] Jackson, S.E., Steptoe, A., Wardle, J. (2015). The Influence of Partner's Behavior on Health Behavior Change: The English Longitudinal Study of Ageing. JAMA Internal Medicine, 175(3):385-392.

[28] Adams, M.A., Sallis, J.F., Norman, G.J., Hovell, M.F., Hekler, E.B., and Perata, E. (2013). An Adaptive Physical Activity Intervention for Overweight Adults: A Randomized Controlled Trial. PloS One, 8(12), e82901. https://doi.org/10.1371/journal.pone.0082901.

[29] Adams, M.A., Hurley, J.C., Todd, M., Bhuiyan, N., Jarrett, C.L., Tucker, W.J., Hollingshead, K.E., and Angadi, S.S. (2017). Adaptive Goal Setting and Financial Incentives. A 2 × 2 Factorial Randomized Controlled Trial to Increase Adults' Physical Activity. BMC Public Health, 17(1), 286. https://doi.org/10.1186/s12889-017-4197-8.

[30] Conn, V.S., Hafdahl, A.R., and Mehr, D.R. (2011). Interventions to Increase Physical Activity Among Healthy Adults: Meta-Analysis of Outcomes. The American Journal of Public Health, 101(4), 751-758. https://doi.org/10.2105/AJPH.2010.194381.

[31] Müller-Riemenschneider, F., Reinhold, T., Nocon, M., and Willich, S.N. (2008). Long-Term Effectiveness of Interventions Promoting Physical Activity. A Systematic Review. Preventive Medicine, 47(4), 354-368. https://doi.org/10.1016/j.ypmed.2008.07.006.

[32] Müller-Riemenschneider, F., Reinhold, T., and Willich, S.N. (2009). Cost-Effectiveness of Interventions Promoting Physical Activity. The British Journal of Sports Medicine, 43(1), 70-76. https://doi.org/10.1136/bjsm.2008.053728.

[33] Rand Corporation (2015). Incentives for Workplace Wellness Programs. They Increase Employee Participation, but Building a Better Program is Almost as Effective. RB-9842-DOL.

[34] Seaverson, E.L.D., Grossnieier, J., Miller, T.M., Anderson, D.R. (2009). The Role of Incentive Design, Incentive Value, Communications Strategy, and Worksite Culture on Health Risk Assessment Participation. Journal of Health Promotion, 23(5), 343-352.

[35] Hooker, S.A., Wooldridge, J.S., Ross, K.M., Masters, K.S. (2018). Do Monetary Incentives Increase Fitness Center Utilization? It Depends. American Journal of Health Promotion, 32(3), 606-612. doi:10.1177/0890117116689321.

[36] Pope, L., Harvey, J. (2014). The Efficacy of Incentives to Motivate Continued Fitness-Center Attendance in College First-Year Students: A Randomized Controlled Trial. Journal of American College Health, 62(2), 81-90. doi:10.1080/07448481.2013.847840.

[37] Jeffery, R.W. (2013). Financial Incentives and Weight Control. Preventative Medicine, 55 (suppl), S61-S67.

[38] Beatty, T.K.M., Katare, B. (2018). Low-Cost Approaches to Increasing Gym Attendance. Journal of Health Economics, 61, 63-76. doi:10.1016/j.jhealeco.2018.05.006.

[39] Abdelsalam, A.E., Said, R.M. (2018). Effect of Health Education Program on Knowledge, Lifestyle and Health Profile of Prediabetic Employees in Zagazig University. JHIPH, 48(1), 1-8.

[40] Prestwich, A., Perugini, M. and Hurling, R. (2009). Can the Effects of Implementation Intentions on Exercise be Enhanced Using Text Messages? Psychol Health, 24(6), 677-687.

[41] Turner-McGrievy, G., and Tate, D. (2011). Tweets, Apps, and Pods: Results of the 6-Month Mobile Pounds Off Digitally (Mobile POD) Randomized Weight-Loss Intervention Among Adults. J Med Internet Res, 13(4), e120.

[42] Harrington, J. (2018). Barriers to Implementation of Workplace Wellness Programs in Career Fire Service Organizations: An Exploratory Qualitative Study. American Journal of Health Studies, 33(1), 40-51. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=asn&AN=131270257&site=ehost-live.

[43] Palmer, B., Baumgartner, N., and Constable, S.H. (2001). Exercise Compliance: Does the US Air Force Have Unique Opportunities? HSIAC-RA-2001-007 Review & Analysis.

[44] Wilkinson, W., Kampert, J., Blair, S., Baumgartner, N. and Constable, S. (2000). Physical Fitness and Health: A Comparative Review of the USAF Fitness Program. USAF School of Aerospace Medicine Technical Report SAM-FE-BR-TP-2000-01.

[45] Baumgartner, N. (2013). Air Force Physical Fitness Test and Standards – Science Rationale. Presented at: National Strength and Conditioning Association's Tactical Strength and Conditioning Association Annual Conference 16 Apr 2013, Norfolk, VA.

[46] US Air Force. (2013). Air Force Instruction 36-2905: Fitness Program. Retrieved from https://static.e-publishing.af.mil/production/1/af_a1/publication/afi36-2905/afi36-2905.pdf.

[47] Gagnon, P., Spivok, M., Reilly, T., Mattie, P., and Stockbrugger, B. (2015). The FORCE Fitness Profile – Adding a Measure of Health-Related Fitness to the Canadian Armed Forces Operation Fitness Evaluation. Journal of Strength and Conditioning Research, 29(11S), S192-S198.

[48] Reilly, T. Aerobic Demands of FORCE Evaluation (2016). Ottawa, Ontario: Canadian Forces Morale and Welfare Services, Internal.

[49] Carstairs, G.L., Ham, D.J., Savage, R.J., Best, S.A., Beck, B., Filling, D.C. (2018). A Method for Developing Organisation-Wide Manual Handling Based Physical Employment Standards in a Military Context. Journal of Science and Medicine in Sport, 21(11), 1162-1167. https://doi.org/10.1016.j.jsams.2018.02.008.

[50] Reilly, T. (2007). Fitness Standards for the Royal National Life Boat Institution (RNLI) Lifeboat Crew. Ph.D. Thesis, University of Portsmouth.

[51] Reilly, T., and Tipton, M. (2010). A Sub-Maximal Occupational Aerobic Fitness Test Alternative, When the Use of Heart Rate is Not Appropriate. Work, 36, 333-337.

[52] Talbot, L. Weinstein, L., and Fleg, J. (2009). Army Physical Fitness Scores Predict Coronary Heart Disease in Army National Guard Soldiers. Mil Med, 174: 245-252, 2009.

[53] Reilly, T., and Tipton, M. (2005). Task-Based Standards for Lifeboat Crew: Avoiding Ageism. International Congress Series, Assessment and Promotion of Work Ability, Health and Well-Being of Aging Workers. 1280, 219-223.

[54] Poplin, G.S., Roe, D.J., Peate, W., Harris, R.B., and Burgess J.L. (2014). The Association of Aerobic Fitness with Injuries in the Fire Service. American Journal of Epidemiology, 179(2), 149-155.

[55] Blair, S.N., Kamper, J.B., Kohl III, H.W., Barlow, C.E., Macera, C.A., Paffenbarder Jr, R.S., and Gibbons, L.W. (2016). Influences of Cardiorespiratory Fitness and Other Precursors on Cardiovascular Disease and All-Cause Mortality in Men and Women. Journal of the American Medical Association, 276, 205-210.

[56] Kodama, S., Saito, K., Tanaka, S., Maki, M., Yachi, Y., Asumi, M., Sugawara, A., Totsucka, K., Shimano, H., Ohashi, Y., Yamanda, N., and Sone, H. (2009). Cardiorespiratory Fitness as a Quantitative Predictor of All-Cause Mortality and Cardiovascular Events in Healthy Men and Women: A Meta-Analysis. Journal of the American Medical Association, 301, 2024-2035.

[57] Stevens, J., Cai, J., Evenson, K.R., and Thomas, R. (2002). Fitness and Fatness as Predictors of Mortality from All Causes and from Cardiovascular Disease in Men and Women in the Lipid Research Clinics Study. American Journal of Epidemiology, 156, 832-841.

[58] Tanaka, H., Monahan, K.D., and Seals, D.R. (2001). Age-Predicted Maximal Heart Rate Revisited. Journal of the American College of Cardiology, 37, 153-156.

[59] Lee, D.C., Artero, E.G., Sui, X., and Blair, S.N. (2010). Mortality Trends in the General Population: The Importance of Cardiorespiratory fitness. Journal of Psychopharmacology, 24, 27-35.

[60] Ryan, R.M and Deci, E.L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Wellbeing. American Psychologist, 55, 68-78.

[61] Spivock, M., Blacklock, R., Newton, P., Stockbrugger, B., O'Hearn, K., Niang, A., and Reilly, T. (2012). Establishing a Baseline Physical Employment Standard for the Canadian Forces: Project FORCE. Presented at First Australian Conference on Physical and Physiological Employment Standards; Canberra, Australia.

[62] Stevens, J., Evenson, K.R., Thomas, O., Cai, J., and Thomas, R. (2004). Associations of Fitness and Fatness with Mortality in Russian and American Men in the Lipids Research Clinics Study. International Journal of Obesity and Related Metabolic Disorders, 28, 1463-1470.

[63] Talbot, L., Weinstein, A., and Fleg, J. (2009). Army Physical Fitness Scores Predict Coronary Heart Disease in Army National Guard soldiers. Military Medicine, 174, 245-252.

[64] Lau, D., Douketis, J., Morrison, K., Hramiak, I., Sharma, A., and Ur, E. (2007). Canadian Clinical Practice Guidelines on the Management and Prevention of Obesity in Adults and Children. Canadian Medical Association Journal, 176, 1103-1106.

[65] Laframboise J.L., Walsh, E.S., and Stockbrugger, B. (2018). Establishing the Relationship Between Cardiorespiratory Fitness, and Performance on the FORCE Evaluation, Age, Sex and Anthropometrics Technical Report: Department of National Defence, Assistant Deputy Minister (Science and Technology), Ottawa.

[66] Hopkins, W.G. (2000). Analysis of Validity by Linear Regression (MS Excel Spreadsheet). In A new View of Statistics. sportsci.org: Internet Society for Sport Science, sportsci.org/resource/stats/xvalid.xls.

[67] McArdle, W.D., Katch, F.I., and Katch V.L. (1991). Exercise Physiology: Energy, Nutrition, and Human Performance. Lea and Febiger.

[68] Baumgartner, N., Logan, R.W., Gruse, M.F., Flerlage, E.M., Hale, K.N., and Batterton, K.A. (2016). USAF Battlefield Airmen Occupationally Specific, Operationally Relevant Physical Fitness Tests and Standards: Prototype vs. Incumbent. Medicine and Science in Sports and Exercise, 48(483). doi: 10.1249/01.mss.0000486451.19977.a2.

[69] Palmer, B., Constable, S.H., and Baumgartner, N. (1997). Expanded Air Force physical fitness battery: muscle strength, muscle endurance, and flexibility considered. USAF Office for Prevention and Health Services Assessment and USAF Armstrong Laboratory Workshop Proceedings PR-97-001.

[70] Palmer, B., Baumgartner, N., and Constable, S.H. (2001). Exercise Compliance: Does the US Air Force Have Unique Opportunities? HSIAC-PR-2001-001 Workshop Proceedings Report, accession number ADA399152.

[71] Hale, K.N., Batterton, K.A., Murphy, E., Baumgartner, N., Logan, R.W., Gruse, M.F., and Flerlage, E.M. (2017). Women, Combat, and Statistics: Developing Occupationally-Specific, Operationally-Relevant, and Gender-Neutral Battlefield Airmen Physical Fitness Tests and Standards. Presented at 84th Military Operations Research Symposium 27 Jan 2017.

[72] Orr, R.M., and Pope, R. (2016). Gender Differences in Load Carriage Injuries of Australian Army soldiers. BMC Musculoskeletal Disorders, 17(488), 1-8.

[73] Baumgartner, N., Constable, S., and Schmidt, D. (1999). Service Member Life Cycle Physical Fitness Integration Symposium. USAF School of Aerospace Medicine Proceedings SAM-FP-BR-PC-2000-0007.

[74] Latimer, M. and Baumgartner, N. (2019). US Air Force Physical Fitness Feedback Survey. Internal communication.

[75] Lee, D.C., Pate, R.R., Lavie, C.J., Sui, X., Church, T.S., and Blair, S.N. (2014). Leisure-Time Running Reduces All-Cause and Cardiovascular Mortality Risk. J Amer Coll Cardiol. 64(5), 472-481.

[76] Smith, S.C. and Haslam, D. (2007). Abdominal Obesity, Waist Circumference and Cardiometabolic Risk: Awareness Among Primary Care Physicians, the General Population and Patients at Risk – the Shape of the Nations survey. Current Med Research and Opinions. 23:29-47.

[77] Robbins, A.S., Chao, S.Y., Baumgartner, N., Runyan, C.N., Oordt, M.S., and Fonseca, V.P. (2006). A Low-Intensity Intervention to Prevent Annual Weight Gain in Active Duty Air Force Members. Mil. Med. 171(6), 556-561.

[78] Bond, S. (2016). The New Rules of Engagement. International Journal of Market Research 58 (3). https://doi.org/10.2501/IJMR-2016-024; https://journals.sagepub.com/doi/abs/10.2501/IJMR-2016-024?journalCode=mrea.

[79] Boston Consulting Group. (2014). How millennials are changing the face of marketing forever. Retrieved from http://eventbrite-s3.s3.amazonaws.com/marketing/Millennials_Research/Gen_PR_Final.pdf.

Chapter 5 – BIOLOGICAL LIMITATIONS TO TASK PERFORMANCE AND TRAINABILITY

M.A. Sharp
US Army Research Institute
of Environmental Medicine
UNITED STATES

T.J. Reilly, M. Cao, and H.C. Tingelstad
Department of National Defence
CANADA

M.C. Canino and S.A. Foulis
US Army Research Institute of Environmental Medicine
UNITED STATES

5.1 INTRODUCTION

Many NATO military forces are removing barriers to women serving in combat occupations. Some forces, such as the Canadian Armed Forces, have long experience with women serving in combat jobs. The Israeli Defense Force maintains a light infantry unit (the Caracal) with specific job assignments for women combatants that is separate from the standard infantry role, which is assigned to men only [1]. The propensity for women to enter these occupations is not high. In a survey of active duty U.S. Army Soldiers, only 19% of women reported that, "if possible, I would consider reclassifying to a combat arms job"[1].

In addition to a small number of interested women, these women must also have the physical capability to perform the critical job tasks of a combat arms job. For example, in 2017 the U.S. Air Force had fewer than ten women compete for entry into the 1,440 entry slots for a Battlefield Airmen and none were accepted.[2] The purpose of this chapter is to briefly discuss the physiological differences between men and women, and how these differences affect a woman's capability to meet the minimum acceptable performance standards of combat arms jobs. The sex bias and adverse impact imposed by the use of different types of selection tests will be examined. The effectiveness of physical training in reducing the sex performance gaps will be described. A database comparing male and female performance on physical fitness selection tests and one comparing military-relevant task simulations are included to inform Physical Employment Standards (PES) development, with specific reference to sex bias and adverse impact. These data may also assist in determining the effectiveness of physical training to reduce sex-related performance differences.

5.2 PHYSIOLOGICAL DIFFERENCES THAT AFFECT COMBAT PERFORMANCE

The sex differences in Soldier performance of physically demanding tasks are due to differences in body size, body composition and physical fitness (i.e., muscular strength, muscular endurance, aerobic capacity, and anaerobic power). A brief synopsis of these differences follows; however, for a more in-depth review of sex-based physiological differences see Greeves [2], Epstein et al. [3], Smith et al. [4], and Roberts et al. [5]. Figure 5-1 is a graphic summary of the physical and physiological differences between the average man and woman. To better understand these difference, review the components of fitness (Chapter 3, Figure 3.3).

[1] Presentation by Dr. Tonia Hefner, "Assessment of the Direct Ground Combat Assignment Rule Exception to Policy", 1 Nov 2012).

[2] Dr. Neil Baumgartner, personal communication, May 2018.

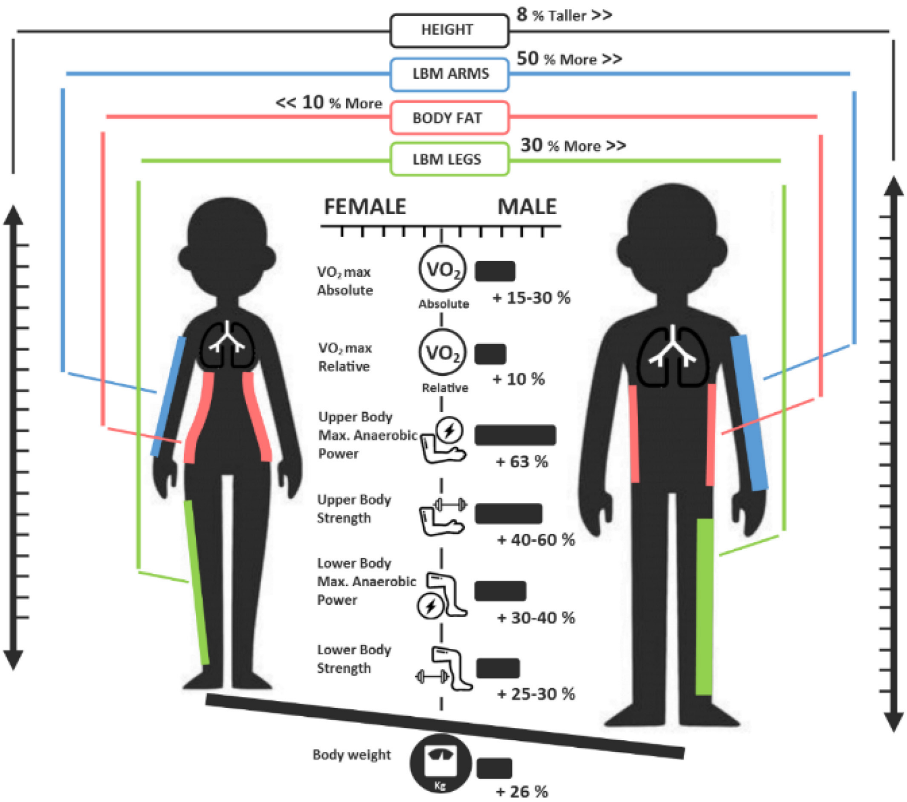


Figure 5-1: Sex Differences in Physical and Physiological Determinants of Physical Performance as Reported in the Literature and Reviewed by Roberts et al. [5] and Reilly et al. [6].

5.2.1 Body Size and Composition

In general, men are of greater stature (8%), greater body mass (26%), and lower-body fat (about 10% less essential fat) than women [7], [8]. In addition, men have more lean body mass than women in their arms (50%) and legs (30%) [9]. Thus, the differences in relative proportions of lean muscle mass and fat mass and the distribution of muscle are different for men and women, which can have profound effects on physical capacities independent of fitness level. It is important to note, however, that although stature is often associated with physical performance, the use of body dimensions is not a legally defensible means of setting PES [10]. Body size is of particular importance when moving external loads such as during manual materials handling and load carriage, two of the most frequently occurring physically demanding tasks of Soldiers [11], [12].

5.2.2 Muscular Strength, Muscular Endurance and Anaerobic Capacity

Women have approximately 70% – 75% of the lower-body strength and 40% – 60% of the upper-body strength of men, with the ratio being influenced by the metric used, age and state of training [13]. Although the average man is stronger than the average woman, there is an overlap in strength such that the strongest women are as strong as, or stronger than, the weakest men [14], [15]. Given smaller body size and lesser strength to move the same object, the average woman must use a greater percentage of her capacity than the average man. Unless she has a higher level of submaximal muscular endurance, she is likely to become fatigued more quickly. To adequately perform a task requiring repetitive heavy lifting, service members must be strong enough that the given absolute workload is a sustainable, submaximal proportion of their maximal lifting strength.

Anaerobic power is important for high-intensity tasks that are performed in an emergency, or while under fire, such as sprinting while wearing a heavy load or quickly dragging a casualty from immediate danger [16]. Values for upper- and lower-body maximal anaerobic power in women are reported to be 37% [17] and 60% – 70% [18] respectively, of those reported for men, which will impact the capability of women to successfully complete tasks requiring rapid application of maximal force.

5.2.3 Cardiorespiratory Endurance/Aerobic Fitness

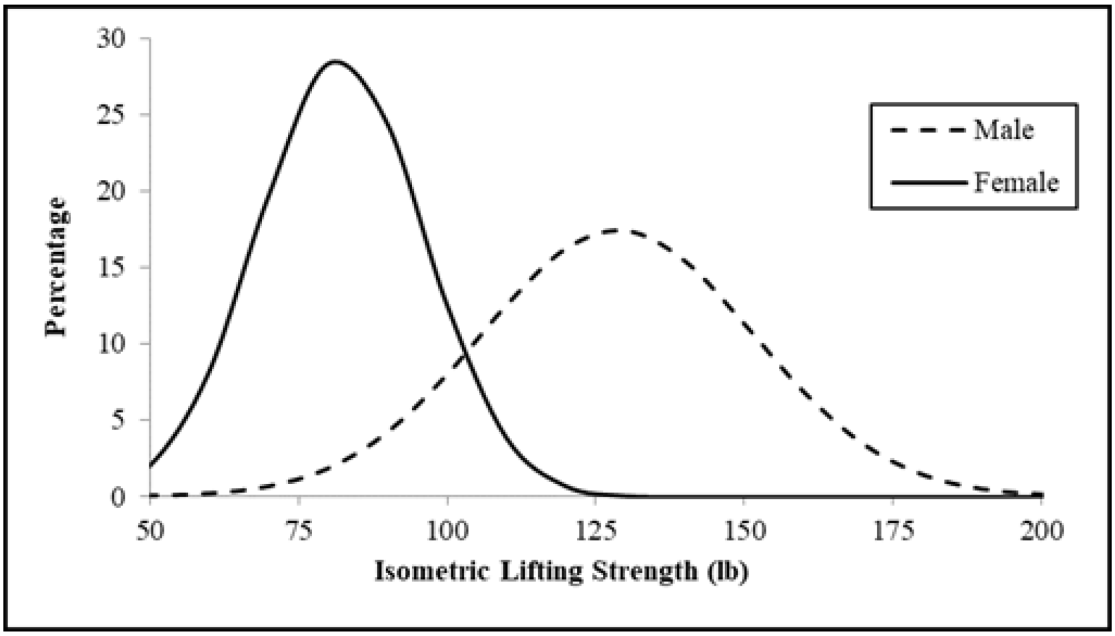
In general, moderately trained women have VO_2max levels that are 15% – 30% lower than moderately trained men in absolute terms ($L \cdot min^{-1}$) and about 10% lower in relative terms ($mL \cdot kg^{-1} \cdot min^{-1}$) [19]. Clearly, the sex differences discussed above for strength and endurance are important to performance of sustained tasks such as repetitive lifting or extended load carriage. The impact of a high VO_2max on repetitive performance of physically demanding tasks suggests that the use of minimum required standards for VO_2max may be an acceptable assessment metric for PES, if the aerobic requirements of the job are accurately defined.

5.3 SEX DIFFERENCES IN SOLDIER TASK PERFORMANCE

Perhaps more important than the physical and physiological differences between men and women are the effects these differences may have on Soldier performance. Figure 5-2 illustrates the comparison of a physical selection test assessing muscle strength and a representative military task test that requires strength. The physical selection test is an isometric lifting strength test [20]. The representative military task test, called the field artillery ammunition loading task, involves loading 45 kg rounds onto an ammunition supply vehicle for distribution to field artillery pieces [21]. The sex distribution is illustrated by a curve for each sex. The areas of overlap indicate the percentage of the measured population of women who performed as well as the measured population of men. Although the overlap is not large, there are women who are as physically fit and who perform as well as some men on a test of basic ability (i.e., muscular strength) and on a test of a representative military task (i.e., ammunition loading task). Figure 5-3 illustrates similar distributions for an aerobic physical selection test (20 m shuttle run test or Beep Test) and performance on an aerobically demanding representative military task (tank ammunition loading task). Again, both show similar distributions and overlapping areas. In both cases, the correlation between the selection test and job task is high.

5.4 FAIRNESS OF PES TESTS

Hodgdon and Jackson [22] define unfairness as "members of a protected group (women) obtain lower scores on a pre-employment test than members of another group (men), but the difference in scores is not reflected in the differences in the criterion of job performance." These authors suggested that a representative military task test is superior to a physical selection test when it comes to test fairness and sex bias, because sex differences in performance may be smaller and representative military task testing is more readily accepted by the applicants. Others have suggested that there is no difference [23], [24]. Examination of the female to male ratios for performance on physical selection tests and representative military task tests does not support a large difference between the test types [25]. In Figure 5-2 and Figure 5-3, the women's scores are distributed similarly for both types of tests (physical selection tests and representative military tasks), which supports the use of either to assess Soldier performance. Courtright et al. [23] suggest that the system of physical selection tests used for PES testing must be considered as a whole, rather than examining each test item individually when comparing the two types of tests. Their meta-analysis found greater sex differences for a job simulation test (i.e., representative military task simulation) than for a single basic ability test (i.e., selection test) that only measures a single component. When a system of basic ability tests (i.e., a test battery containing multiple components of fitness, was compared to a job simulation test), the sex differences between the two types of tests were comparable. This is supported by the work of Gebhardt and Baker [24], [26], who recommend considerations such as validity, reliability, cost and worker acceptance be considered when making the decision of test type.



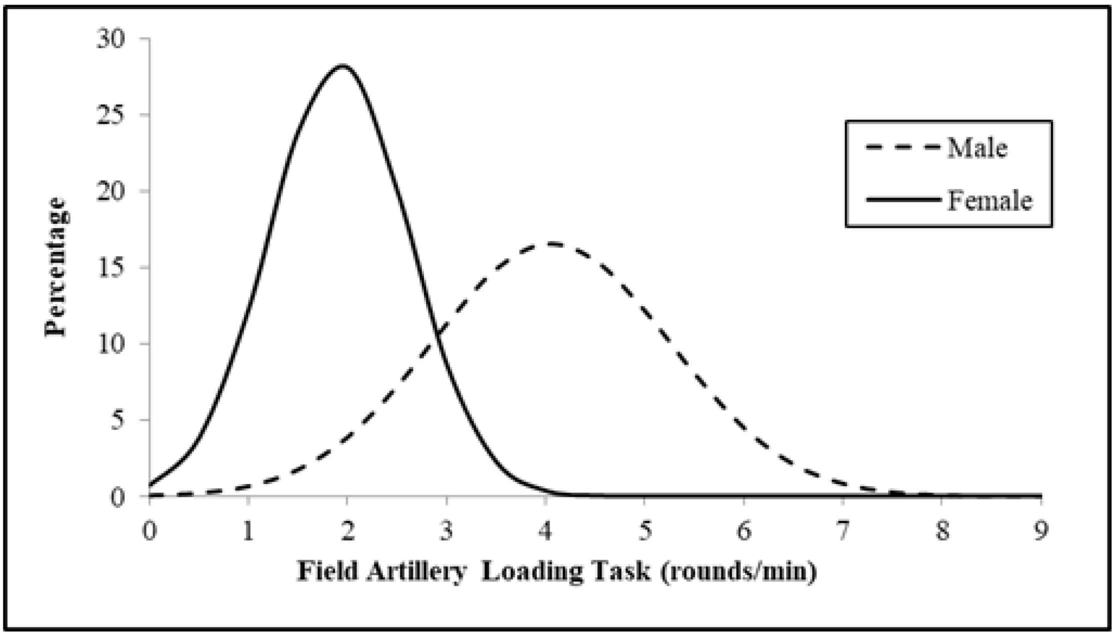


Figure 5-2: Male-Female Comparison for the Isometric Upright Pull Physical Selection Test (Upper Graph) and for the Field Artillery Ammunition Loading Task (Lower Graph).

Note: The correlation between isometric lifting strength and the field artillery ammunition loading task is 0.77.

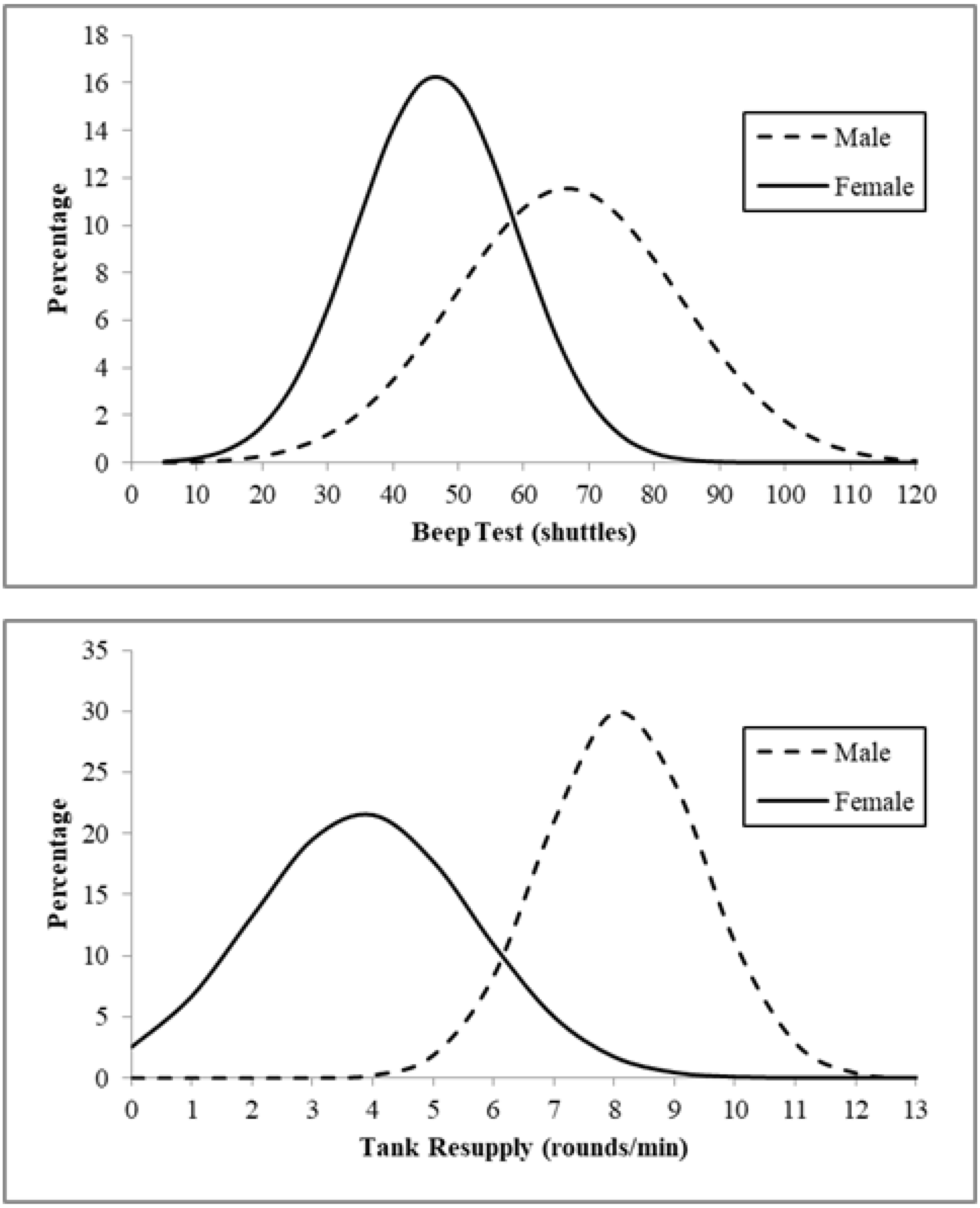


Figure 5-3: Sex Distribution of Scores on the Beep Test (Top Graph) Selection Test and Tank Resupply (Bottom Graph) Representative Military Task.

Note: The correlation between the beep test and the tank resupply task is 0.65.

The predictive relationship between isometric lifting strength and the field artillery ammunition loading task is illustrated in a scatterplot (Figure 5-4) for men and women [21]. The predictive capacity $r^2 = 0.57$ is reduced to an equivalent extent in single gender analysis (r^2 men = 0.23, r^2 women = 0.24). The slope and intercept of the gender-specific equations were not statistically different. In this case, there is no sex bias and the isometric lifting strength test can be used to predict performance of men and women on the field artillery ammunition loading task. The required passing score on the field artillery ammunition loading task is shown as a dotted horizontal line. Those who would be selected for the job are shown in the far right section (Prepared), those who would fail the test are in the left hand side (Unlikely to Succeed). The graph also displays the false passing (in the Prepared section but below the dashed line) and false failing Soliders (in the Unprepared section, but at or above the dashed line). There is also an area marked 'trainable'. While there is no way to easily deteremine an individual's trainability, it is estimated that those scoring below the standard in this area might be capable of training to pass the test. This is a simplified example which included only strength as a predictive measure. A multivariate regresson analysis, which included other essential physical abilities, would increase the predictive capacity, while reducing the error of prediction.

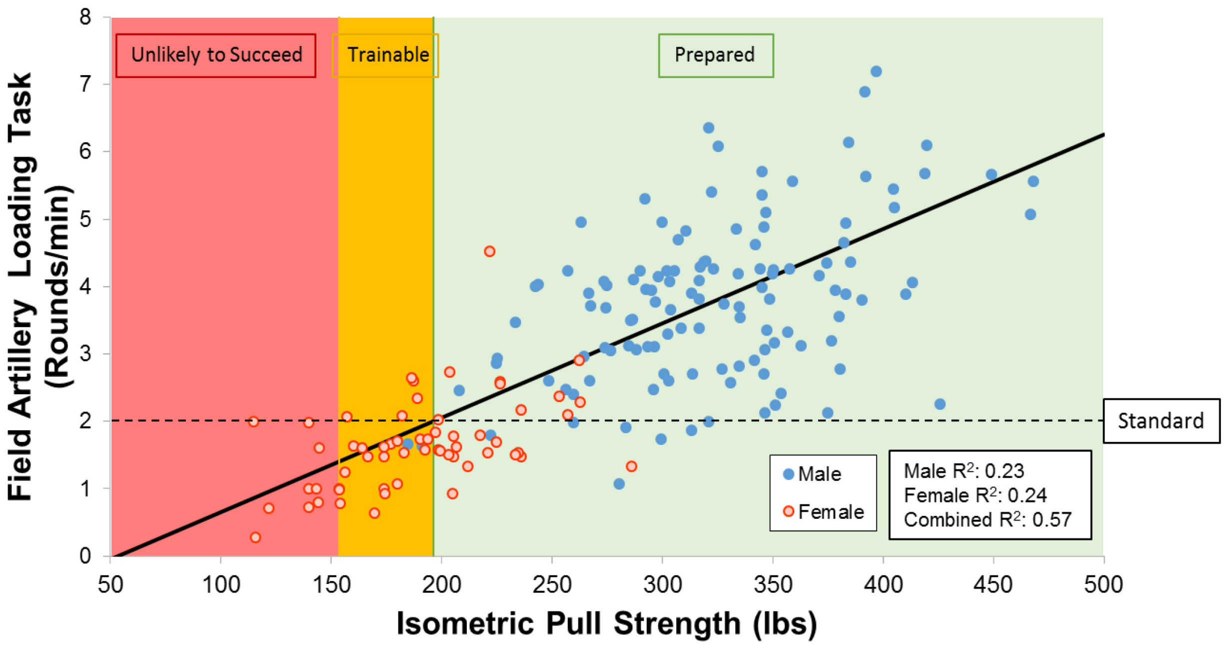


Figure 5-4: Predictive Relationship Between Isometric Lifting Strength (Selection Test) and the Field Artillery Ammunition Loading Task (Representative Military Task).

5.5 PHYSICAL TRAINING TO IMPROVE PERFORMANCE IN PHYSICALLY DEMANDING JOBS

Appropriate physical training has been used to improve performance of physically demanding job tasks in both men and women [27]. Aerobic training has been shown to result in improved performance in tasks with a high aerobic demand in Soldiers [28], [29], [30]. A number of studies have shown that combined resistance and aerobic training programs can be used to improve the performance of women on physically demanding tasks such as lifting and load carriage [31], [32], [33], [34]. Kraemer et al. [33] reported that a six-month combined strength and endurance training program improved the repetitive lifting and load carriage performance of women such that there was no sex difference when compared to untrained men. Implementing high-intensity interval training may provide a time efficient means to improve a Soldier's aerobic fitness [34]. This is supported by Drain et al. [35], who reported a periodized, lower volume, higher intensity training program produced greater gains in women for strength-related tests and Soldiering tasks, as compared to the standard recruit physical training program. The women in the high-intensity training group improved push-up and box lift-and-place task performance twice as much as the control group women. A review of physical training programs used to improve women's performance of military occupational tasks concluded that a minimum of six months training using exercises specific to job tasks, and load carriage specific training, with an emphasis on upper-body strength and power were necessary to prepare women for combat occupations [27].

In Figure 5-4, there is a central section labelled 'Trainable'. These individuals tend to cluster close to the pass/fail line making it difficult to deny an applicant the job. With an optimal training program, manual materials handling performance in women can be increased by 30% or more depending on the type of task, initial fitness of the Soldiers, and specific training procedures utilized [31], [36], [14]. The decision to accept applicants in the trainable group is as much an administrative decision as a scientific one. It is based on risk management and the number of available slots and applicants. In order to perform the same job tasks, women must often work at a higher percentage of their maximal capacity than men. While women may be able to train to reach an acceptable level of performance, not all will have the necessary reserve to respond to emergencies when a higher intensity of performance is needed [5]. In addition, it is unclear the degree to which increased physical fitness resulting from training will be protective of injuries over time. However, fitness is related to musculoskeletal injury in military recruits undergoing basic training, regardless of sex [37].

In a recent randomized controlled training study, Gumeiniak et al. [38] demonstrated the effects of both test practice and physical training on reducing the sex bias in PES assessment. Using the Canadian Wildland Firefighter Fitness Test circuit as the criterion measure, only 11% of women were able to pass the test to standard on the first try, as compared to 73% of men. Volunteers were divided into three cohorts: a physical training cohort, a circuit training cohort and a control group. At the end of 5 weeks physical training, 5 weeks of performing the circuit once each week or doing nothing (control), 80% of the training cohort females passed, 72% of the circuit-only females passed, but only 26% of the control group passed. These researchers concluded physical training and adequate test familiarization significantly reduced the sex bias in the test [38]. Courtright et al. [23] also examined training as a moderator of sex differences in physical performance as part of their meta-analysis. They concluded that training is a successful moderator of sex bias if it enables women to meet the minimum acceptable performance standards to pass a PES assessment and perform the job. Physical training can assist women in gaining access to the job, but must be continued throughout their employment to maintain an acceptable level of performance of physically demanding job tasks [5].

5.6 COMPENDIUM OF PHYSICAL ABILITY TESTS AND MILITARY-RELEVANT PHYSICALLY DEMANDING TASKS IN MILITARY PERSONNEL

One of the goals of this NATO group was to share the data available to compare male and female performance. Two databases were constructed based primarily on performance in male and female military personnel including Army, Navy and Air Force trainees and incumbents. Some data that were available in the literature from civilian emergency services personnel (i.e., firefighters and police) as well as some physically demanding occupations were also included. Existing performance data from peer-reviewed and technical publications, as well as some unpublished data were gathered from the following countries: Australia, Canada, Denmark, France, Germany, Israel, Netherlands, New Zealand, Norway, the United Kingdom, and the United States. One database consists of physical selection tests, also known as physical ability tests. These include measures such as handgrip strength, beep test or vertical jump performance and were grouped according to test description. The physical selection tests were performed to maximal effort in order to qualify for inclusion into the compendium. The second database contains representative military tasks performed by one person, or a few team tasks that were pro-rated for one person. These tasks are physically demanding and were also grouped according to task description into a series of tables (movement tasks, climbing tasks, etc.), they are also depicted in figures within the electronic database (some examples will be provided here). The representative military tasks were typically performed at high intensity or at least at an intensity for minimally acceptable military performance.

Sex-specific means and standard deviations were essential for a study to be included in the databases. The purpose of these databases is to assist in selecting physical selection tests and representative military task tests with the least sex bias, or to inform the researcher of the bias that can be expected when using similar tests. These weighted averages and standard deviation values were used to create probability density curves, to demonstrate the percentage of overlap between male and female performance. The percentage of overlap was calculated as the area of intersection between the male-female curves which has the specific aim to illustrate the potential to combine male and female data in any statistical analysis to reduce sex bias. For male and trained female samples with available training data, the percentage of overlap pre- and post-training was calculated to highlight the effects of training on reducing sex-related performance differences. For select timed Physical Fitness Tests (PFTs) and tasks, shorter durations represent better performance, which is why some female probability density curves are right of the males.

5.6.1 Physical Selection Test Data

Table 5-1 contains a summary of the physical selection tests included in the database. Raw data were classified and sample size weighted means and standard deviations were developed for each test. Probability density curves were created to assess the percentage of overlap between male and female performance. Where female training data were available, the change in percentage overlap is presented to illustrate the potential reduction in sex-related performance differences following physical training. As shown in Figure 5-5, the physical ability test with the least sex disparity was sit-ups, while bench press had the greatest sex disparity.

Table 5-1: Sex-Specific Physical Selection Test Standardized Means (N = 158,542 Male Data Points, 43,460 Female Data Points).

Physical Fitness Test	Metric	Male Mean (SD)	N Male	N Female	Female Mean (SD)	F:M Overlap	Cohen's D (Non-Training)
Aerobic Fitness							
VO_{2max}*	mL/kg/min	46.5 (10.7)	2831	1693	36.6 (10.2)	59.9%	0.9
VO_{2max}	L/min	4.0 (0.5)	161	80	2.4 (0.3)	8.8%	3.3
3.2 km run*	min	14.9 (1.4)	6249	1476	18.0 (2.0)	40.2%	1.8
PARE **	sec	225.0 (42)	21	27	341.0 (97)	38.2%	1.6
2.4 km run*	min	13.3 (1.3)	4153 1	6194	15.3 (1.5)	34.2%	1.4
3 km run	min	13.6 (1)	3713 5	5236	15.8 (1)	28.4%	2.1
Anaerobic Fitness							
Arm ergometer	rev	252.3 (34)	771	273	189.7 (34.6)	35.4%	1.8
100 m sprint	sec	14.3 (1.4)	160	57	17.0 (2.2)	18.0%	1.4
40 m dash	sec	5.6 (0.4)	82	34	7.0 (0.5)	15.6%	3.1
300 m sprint*	sec	47.5 (4.2)	3185	883	58.4 (5.4)	15.3%	2.3

Physical Fitness Test	Metric	Male Mean (SD)	N Male	N Female	Female Mean (SD)	F:M Overlap	Cohen's D (Non-Training)
Muscular Strength							
Leg press	kg	314.0 (82.0)	26426	11541	214.9 (62.7)	53.7%	1.4
Static lift	kg	126.4 (23.8)	2240	1849	80.1 (34.5)	33.3%	1.6
Back extension	kg	93.4 (16.1)	298	153	70.3 (10.4)	37.6%	1.7
Grip strength*	kg	48.3 (15.8)	2640	1586	30.9 (10.7)	23.1%	1.3
Combined grip	kg	103.4 (18.1)	17376	2217	64.8 (21.5)	21.5%	1.9
Squat*	kg	99.0 (18.3)	668	621	69.7 (14.0)	24.0%	1.8
Isometric arm curl	kg	44.4 (8.1)	163	43	27.6 (4.4)	15.0%	2.6
Bicep curl	kg	41.6 (6.7)	139	72	24.2 (4.6)	15.2%	3.0
Bench press*	kg	93.9 (22.3)	2837	389	44.2 (8.6)	8.8%	2.9
Incremental lift	kg	61 (12.4)	18536	7172	30.2 (5.9)	49.5%	3.2
Muscular Endurance							
Sit-ups 1 min.*	#	34.9 (18.1)	21280	4014	36.4 (18.1)	93.4%	0.1
Sit-ups 2 min.*	#	58.6 (13.1)	5522	208	50.2 (15.3)	73.9%	0.6
Flexed arm hang	sec	53.0 (19.3)	561	288	28.1 (16.3)	65.2%	1.0
Push-ups 1 min.	#	48.6 (30.3)	817	314	28.8 (19.3)	49.1%	0.8
Push-ups 2 min*	#	42.9 (11.9)	11640	3188	19.4 (10.4)	15.2%	2.1
Pull-ups*	#	7.6 (4.1)	3227	945	0.9 (2.1)	26.9%	2.1
Muscular Power							
Vertical jump	cm	51.2 (8.8)	166	79	37.3 (4.6)	30.8%	2.0
Standing long jump	m	2.1 (0.2)	501	424	1.6 (0.2)	26.6%	2.4
Counter movement vertical jump*	W	1262.0 (159)	7	6	913.0 (92)	16.6%	2.7
Wingate	W/kg	10.4 (1.5)	741	226	6.4 (1.2)	13.2%	2.9
Flexibility							
Sit and reach	cm	23.7 (8.1)	282	80	24.7 (11.6)	91.6%	0.1

* Denotes training data is reported.

** Physical Ability Requirement Evaluation for the Royal Canadian Mounted Police, consisting of: Obstacle Course, Push/Pull Station, and Torso Bag Carry.

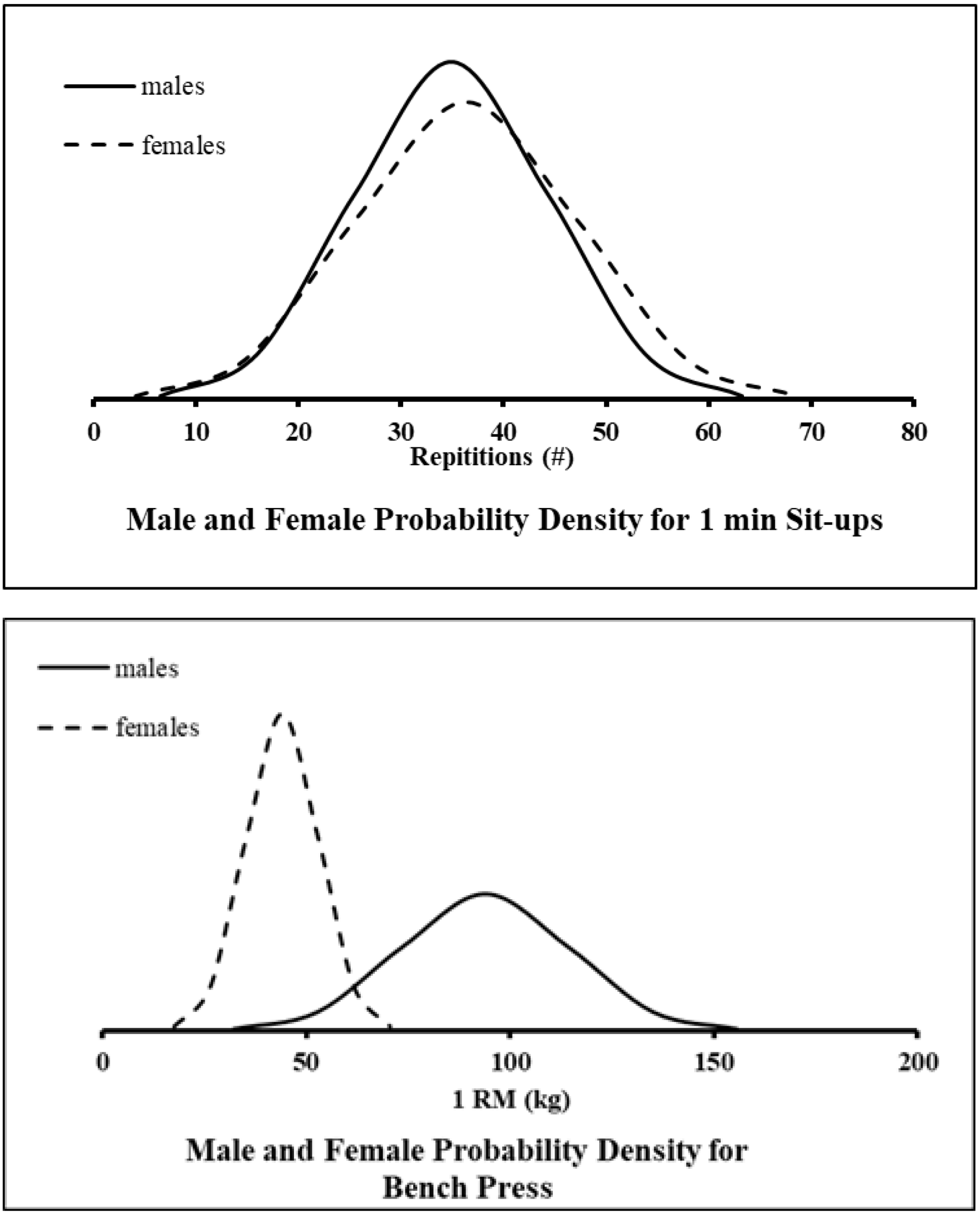


Figure 5-5: Normalized Population Histograms Illustrating Sex Bias for Sit-Ups (Upper Graph), Which had the Smallest Sex Disparity and Bench Press (Lower Graph), Which had the Largest Sex Disparity.

Pre- and post-training data were found for women performing 12 of the physical selection tests. Examination of the sex disparity before and after training revealed which physical selection tests might most benefit from physical training. Training for bench press, pull-ups, VO_{2max}, and upright pull improved female performance by 12%, 22%, 35%, and 23%, respectively. This translated into narrowing the gap between male and female mean performance by 1%, 4%, 5%, and 10%, respectively.

Bench press performance (Figure 5-6(A) and (B)) was determined by 1 repetition maximum (1RM) from military, police and firefighters in American, Norwegian and Canadian populations. Figure 5-6(A) shows the overlap in male and female performance is minimal (8.8%) with a Cohen's d of 2.9. Training data (Figure 5-6(B)) shows improvement of up to 10% overlap, an increase of 1.2 percentage units, but the sample size for females is very small (n = 19).

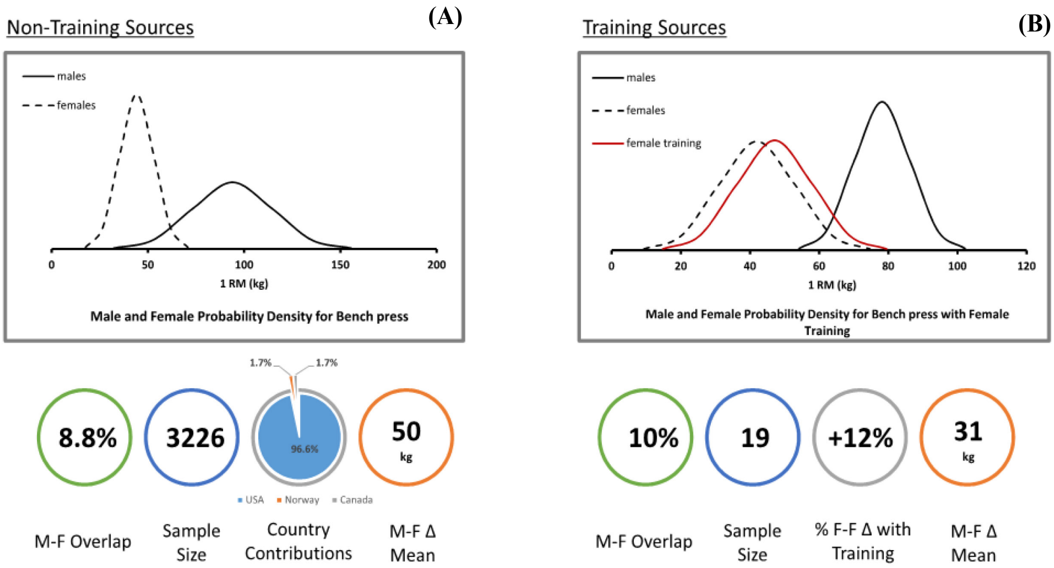


Figure 5-6: (A) Probability Density Curves for Bench Press Strength of Men and Women; (B) Effect of Training on Probability Density Curve for Bench Press Strength of Men and Women.

Looking at pull-ups (Figure 5-7(A) and Figure 5-7(B)), the data reveal a relatively small amount of overlap in performance (27%), and that the male-female performance overlap increases 4.2 percentage units with physical training of females (N = 641 females).

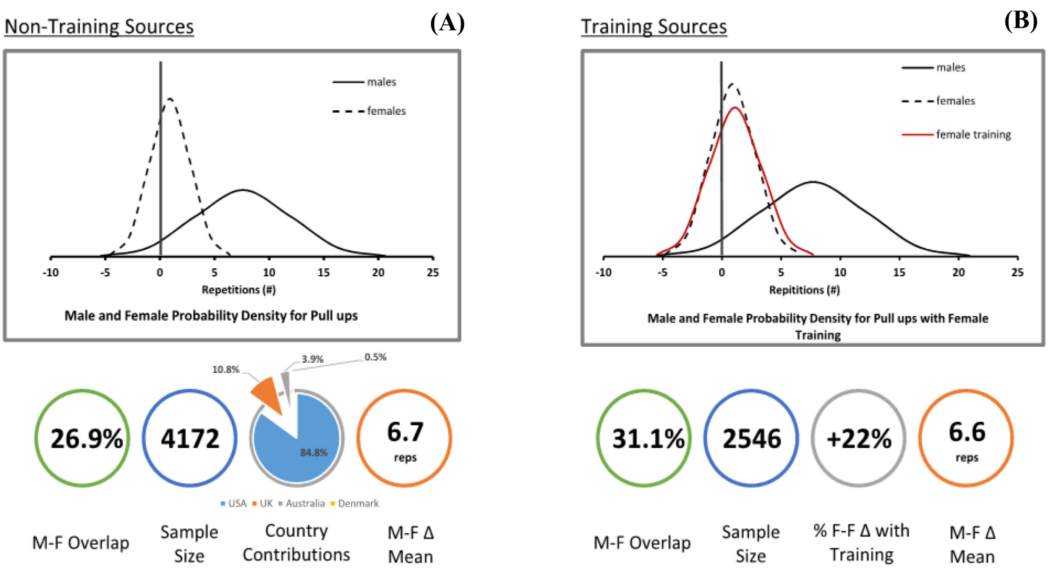


Figure 5-7: (A) Probability Density Curves for Pull-Ups in Men and Women; (B) Effect of Training on Probability Density Curve for Pull-Ups in Men and Women.

The VO$_2$max data (Figure 5-8(A) and Figure 5-8(B)) show that female performance showed improvement with training of 5 percentage points to 65% overlap, which reduced the difference between sexes from a pre-training difference of 9.8 mL•kg^{-1}•min^{-1} to a post-training difference of 5.8 ml•kg^{-1}•min^{-1}.

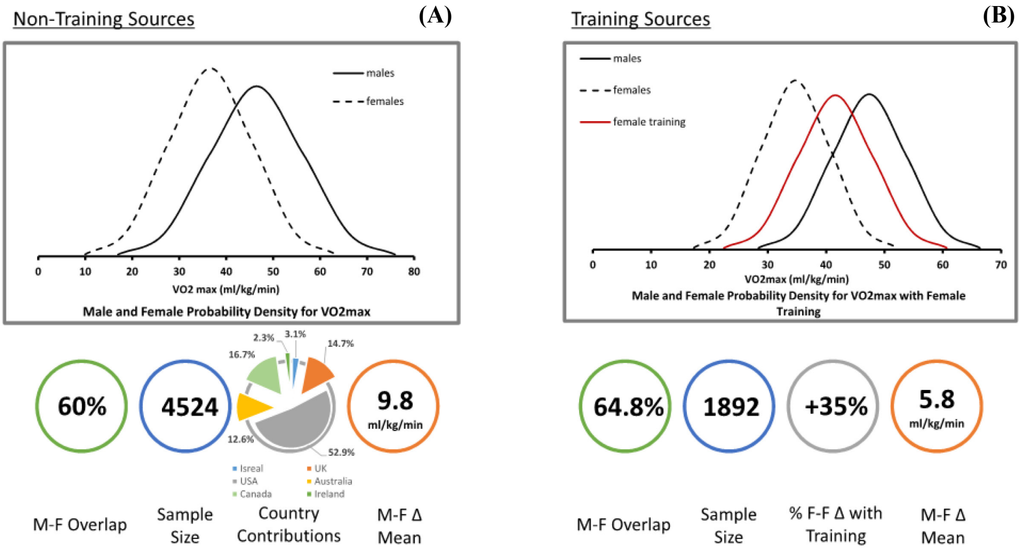


Figure 5-8: (A) Probability Density Curves for VO$_2$max Expressed Relative to Body Mass (ml•kg^{-1}•min^{-1}) in Men and Women; (B) Effect of Training on Probability Density Curve for VO$_2$max Expressed Relative to Body Mass (ml•kg^{-1}•min^{-1}) in Men and Women.

Following training, there was an improvement of 10 percentage points on the maximum weight (force) of the isometric upright pull (Figure 5-9(A) and Figure 5-9(B)), achieving 42% overlap. These data were obtained from 592 females before and after 8 weeks of U.S. Army Basic Combat Training [31], [39].

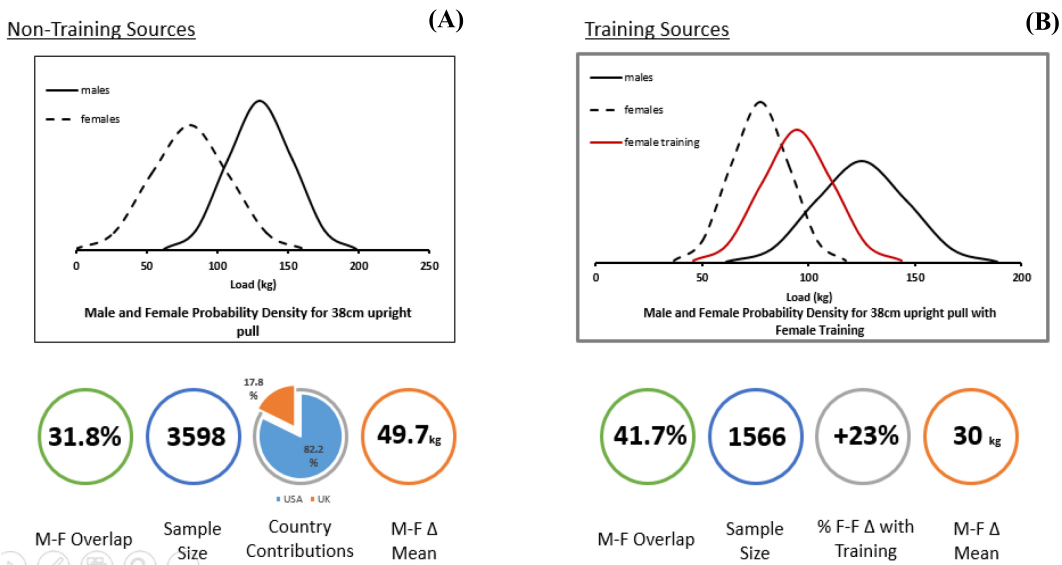


Figure 5-9(A): Probability Density Curves for Upright Pull (Described as an Isometric Upward Pull at 38 cm from the Ground) in Men and Women; (B) Effect of Training on Probability Density Curve for Upright Pull in Men and Women.

These findings should be viewed with caution because the quality of the training programs was not evaluated, nor was the initial fitness level of the participants taken into account. Despite these limitations, these comparisons provide a valuable resource for PES researchers to develop a protocol while minimizing sex bias. The ability of a physical selection test to predict military-relevant task performance is the most important factor; however, if physical selection tests with more overlap are chosen, this may reduce sex bias when sex-free PES are enforced. Additionally, examination of potential training effects in physical selection test performance informs the researcher on which physical selection tests have the most potential for improvement in female populations. This information can also be used to reduce sex bias in PES assessment.

5.6.2 Task Simulation Data

The task simulation database is extremely varied and difficult to neatly summarize. By way of example, a sample table of some of the repetitive lift and carry task simulations with performance standards is shown in Table 5-2. Similar tables are available on request for all the task categories, which include climbing, combat movement, digging, dragging/pulling, manual materials handling, load carriage, and obstacle course / other multi-tasks. The MMH task categories contained the greatest number of tasks. The tasks were classified as a single maximal effort, which included lifts of loaded boxes and power bags. Often the tasks had a maximum allowable weight, which ranged from as low as 40 kg to as high as 100 kg. The F:M ratio ranged from as low as 10% for the U.S. Navy 1RM box lifting task to as high as 71% for the UK power bag lifting task. The majority of the F:M ratios were in the range of 50 – 70%. There was only one repetitive lifting task that did not also involve carrying, the Canadian Armed Forces sandbag lifting task, which had a F:M ratio of 60%.

There were many variations of continuous lift and carry tasks involving water jugs, power bags and ammunition boxes. The F:M ratio ranged from as low as 41% to as high as 87%. The repetitive lift and carry tasks utilized the same items as the continuous lift and carry tasks, but involved an unloaded walking portion of the task (see Table 5-2). The F:M ratios were very similar to those of continuous carries and ranged from 44% to 89%.

The total load, uniform and distance moved for the load carriage tasks varied greatly across the tasks. Loads ranged from 15 kg to 49 kg and distances ranged from 3.2 km to 12.8 km. The F:M ratio (inverse for timed events) ranged from 64% to 85%. Two reported digging tasks involved moving gravel from one place to another with F:M ratios of 57 and 80%. The fire and manoeuvre tasks involved bounding, running and low-crawling (also referred to as leopard crawling). The F:M ratios were high and ranged from 72% – 98%. The reported obstacle course data were from emergency responders, and the F:M ratios ranged from 73 – 86%. Although it is difficult to draw generalized conclusions about the tasks regarding sex bias, it seems that women perform more comparably to men on lower intensity, prolonged and in some cases multi-element tasks. These data may be used to identify problem areas for sex bias in military tasks and to develop training programs aimed at improving the military performance of women.

Having a large database of simulation data also allows for investigation into how simulation design of a simple task, like stretcher carry, can affect the performance ratio between males and females. The most common simulation of stretcher carry is a jerry can carry or carrying a deadlift barbell. The purpose of a stretcher carry is to transport a wounded Soldier or civilian to a triage point [40], [41]; however, there is currently no consensus on the distance they have to be carried, the load of the wounded Soldier, or the speed of carry. Therefore, the simulation database can be used to explore how differences in load, distance and speed can affect performance of males and females.

Table 5-2: Repetitive Lift and Carry Tasks and the Female to Male Performance Difference (Expressed as a Percentage).

Country	Task Variables	Uniform Load	Metric	Female Mean (SD)	Male Mean (SD)	F:M Ratio
AU Army	2 x 22 kg jerry can carry, 6 x 25m shuttles, set speed		m	204.1 (97.5)	444.3 (198.8)	46%
UK Army	10 kg ammo box, 10-m carry, 1.45-m lift, return, up to 60 min		sec	2311	3574	65%
UK Army	22 kg ammo box, 10-m carry, 1.45-m lift, return, up to 60 min		sec	1048	3578	29%
UK Army	20 kg sandbags, 30-m carry, 1.1-m lift, AMAP* in 10 min		# of sandbags	13.4	17.7	76%
US Navy	34 kg box, 51 m, 2 x 5 min (1 min rest), AMAP carries		Watts	271 (37)	305 (39)	89%
NZ Army	2 x 20 kg jerry can carry, 8 x 25 m, 5-sec rest between shuttles, 4.5 km/hr		Pass/Fail	Na	Na	78%
US Army	16, 18 kg sandbags, carry 10 m	29 kg	Time	3.0 (1.1)	1.7 (0.3)	58%
US Army	30, 45 kg projectiles, carry 5 m, floor to shoulder lift in 15 min	22 kg	rounds/min	1.6 (0.7)	3.8 (1.2)	43%
US Army	Carry 18, 25 kg rounds 5 m, lift to 163 cm	29 kg	rounds/min	3.4 (1.8)	7.6 (1.3)	44%
CAF	20 kg sandbags, carry 50 m, AMAP in 10 min		# of sandbags	9.5 (1.4)	12.1 (2.6)	79%
US Army	25 kg box, carry 5 m, AMAP in 5 min		# of carries	23.7 (4.7/min)	37.2 (7.4/min)	63.71%
US Army	25 kg box, carry 5 m, AMAP in 10 min		# of carries	41.2 (4.1/min)	66.7 (6.7/min)	61.77%
US Army	45 kg box, carry 5 m, AMAP in 5 min		# of carries	17 (3.4/min)	36.6 (7.3/min)	46.45%
US Army	45 kg box, carry 5 m, AMAP in 10 min		# of carries	9.4 (0.9/min)	20.6 (2.1/min)	45.63%
*AMAP = as many as possible						

Figure 5-10(A) through Figure 5-10 (E) depict the normal curves (mean ± SD) for stretcher carry tasks that were simulated by carrying jerry cans. The task performed is the same in each figure; however, the outcome measure (speed, time / distance or load) is different. When distance carried is the outcome measure (Figure 5-10(A)), the F:M performance ratio was 73.4%, whereas when performance was based on maximal load carried (Figure 5-10(B)), the F:M performance ratio was only 60.2%. When carrying 2 x 20 kg jerry cans for time to exhaustion the female to male performance ratio was only 46% (Figure 5-10(C)) or 40.6% (Figure 5-10(D)). When the outcome measure was speed (carrying 2 x 18 kg jerry cans 100 m as fast as possible, [Figure 5-10(E)]), the F:M performance ratio was 62.9%.

The minimum acceptable performance standard (MAPS) selected also exerts a great effect on the relative success of women. Although the F:M performance ratio is 46% on the 2 x 22 kg jerry can carry (Figure 5-10(C)), the pass rate for males and females is 93.6% and 72.6% respectively, with MAPS of ≥ 150 m. On the other hand, the F:M performance ratio on 2 x 18 kg jerry can carry task where performance is based on speed was as high as 62.9%, however, pass rate was low; 37.1% for males, and only 1.8% for females with a MAPS of ≤ 50 sec (GER Army Soldiers). Since litter carriage is a critical and commonly performed task, it is possible this low pass rate is due to unrealistic standards, perhaps based on speeds that may not be grounded in practical relevance (see discussion in Chapter 3).

Similar visualizations, calculations and analysis can be done for all the tasks included in the database, and can be a valuable tool in the development of new physical employment standards, and the tailoring of procedures and constraints to decrease bias towards females. The following figures depict the current ability of the electronic performance comparison compendium to compare task based tests. Figure 5-11 demonstrates that for four sources of loaded marching performance data, all measured by time to completion, 59% of females achieved a pass and 97% of males achieved a pass.

This information can be further broken down by each task by selecting:

MORE INFORMATION

and the various protocols used in each reference are explained, demonstrating pass and fail statistics dependent on each protocol. This breakdown allows the user to identify which protocol best addresses the operational needs of the task while also considering the expected outcomes of performance for each sex (Figure 5-12). These breakdown analyses allow the same normal density curves to be produced for each protocol (described in the task description box) to compare actual performance and observe the overlap of male and female data (Figure 5-13).

5.7 CONCLUSION

Although there are established differences in the physical capabilities of men and women, it is important to recognize that there are some women capable of performing all tasks to standard, just as there are some men who are not [50]. The performance curves for physical fitness tests and for task simulations typically have some overlap. Physical training for women can provide an effective means of reducing sex bias and has been shown to improve job performance in both sexes. Selecting PES assessment measures with the greatest overlap that are also good predictors of job performance will reduce sex bias. The metric and the testing method used to evaluate task performance will also produce differences in female to male performance ratio. The greater the intensity of the task (load, speed, number of repetitions), the greater the impact on women's performance. The two databases provide useful information for researchers aiming to reduce sex bias when developing PES assessments and provide comparison data for similar tasks.

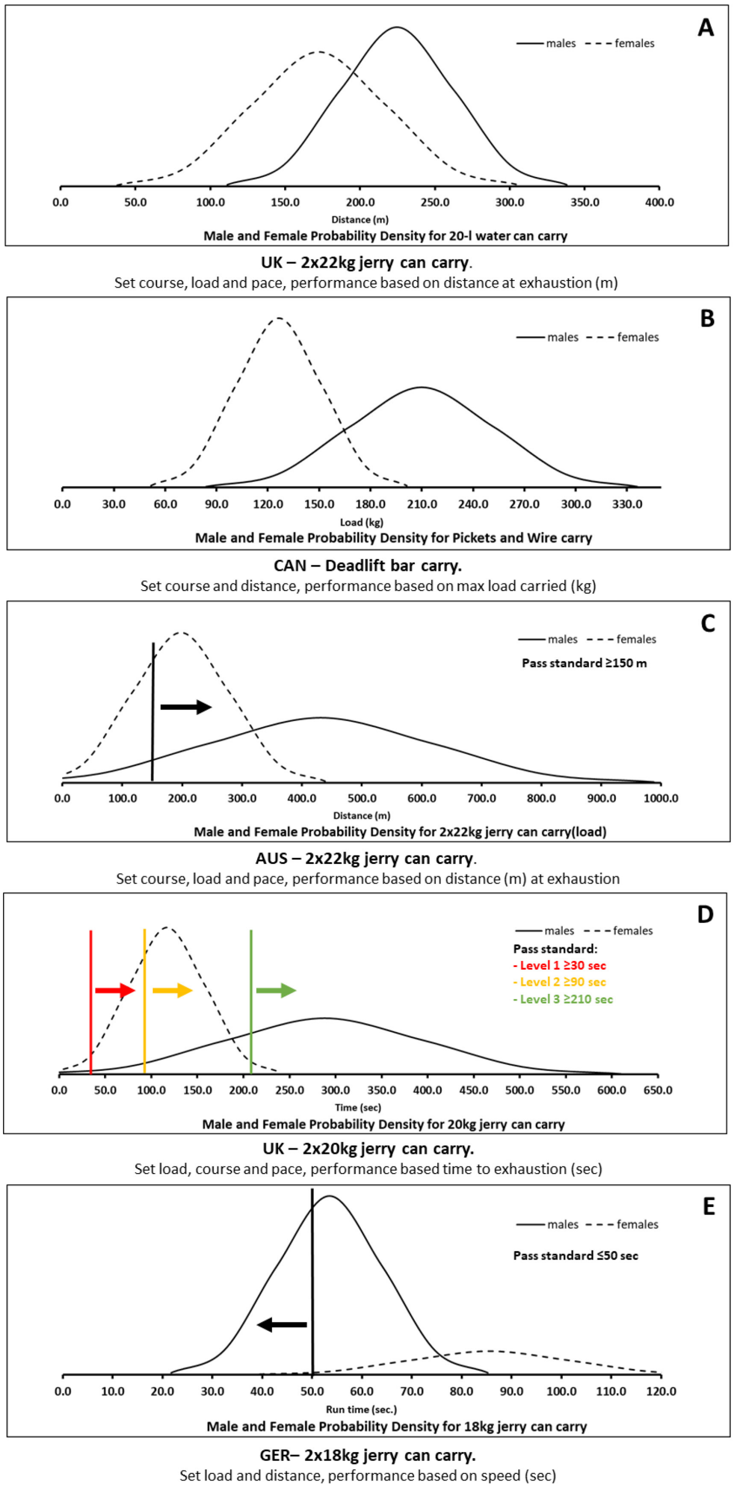


Figure 5-10: Probability Density Curves.

Note: A) Set pace water can carry in men and women measured as distance completed [42], [43], [44], [45]; B) Water can carry in men and women measured as maximum load carried [41]); C) Repeated lift and carry 2 x 22 kg jerry cans to a cadence in men and women for maximum distance possible [46]; D) Carrying two water cans for as long as possible [47]; E) repeated lift and carry of 2 x 18 kg jerry cans for time [48], [49].

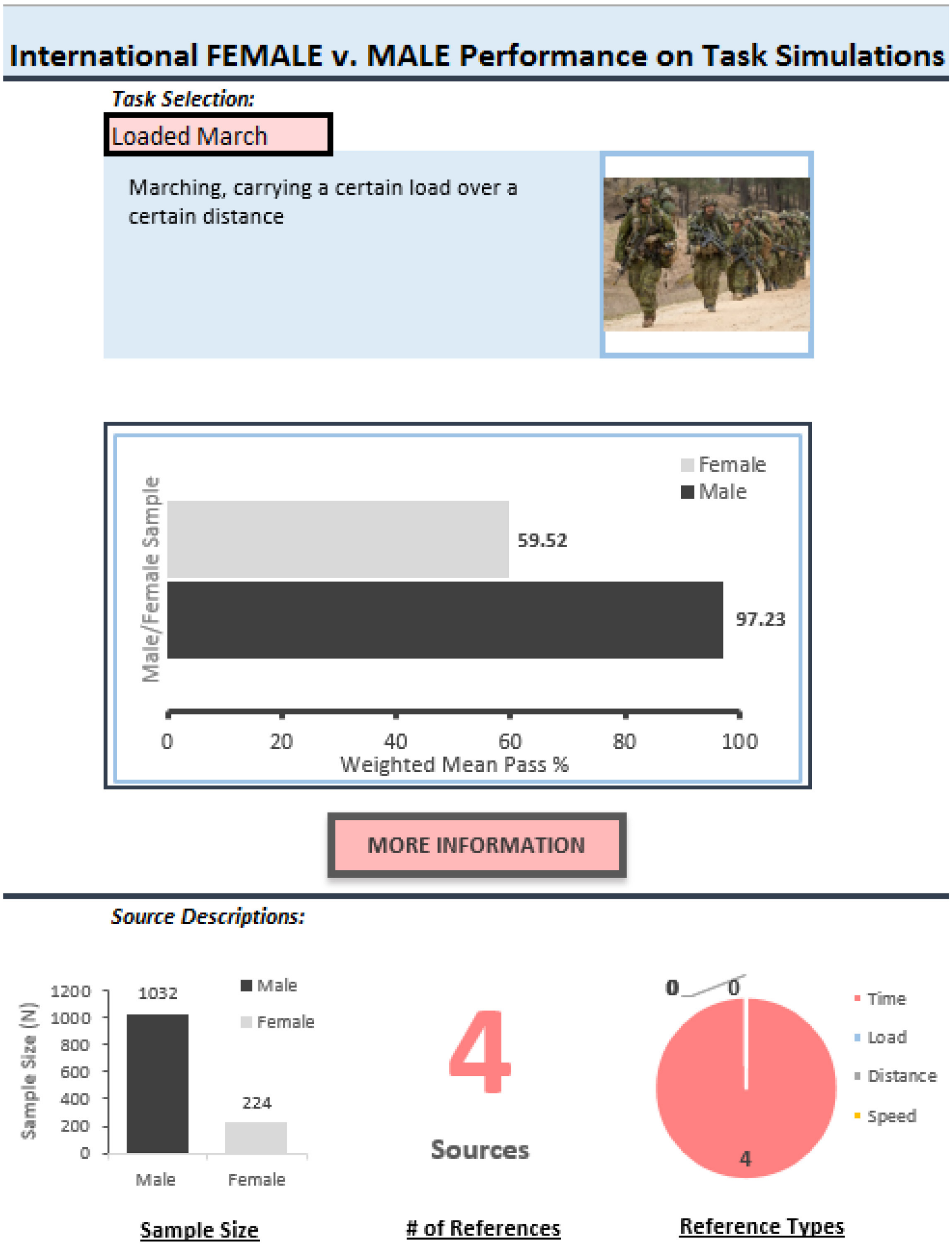


Figure 5-11: The Main Interface of the Electronic Performance Comparison Compendium (EPCC) for the Loaded March Tas

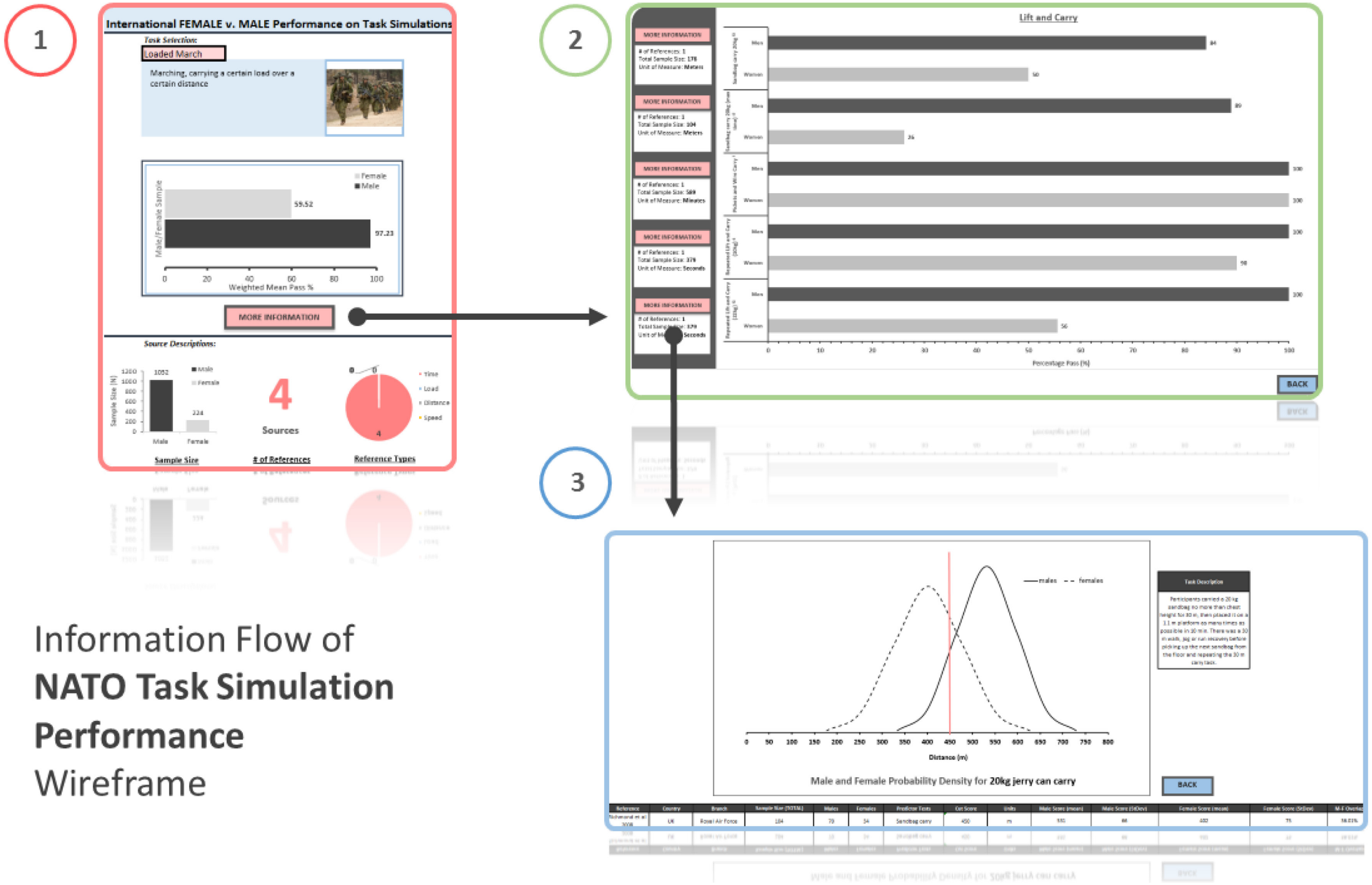


Figure 5-12: The Information Flow of the Electronic Performance Comparison Compendium (EPCC), from the Main Interface, to the Task Breakdown, to the Male – Female Overlap.

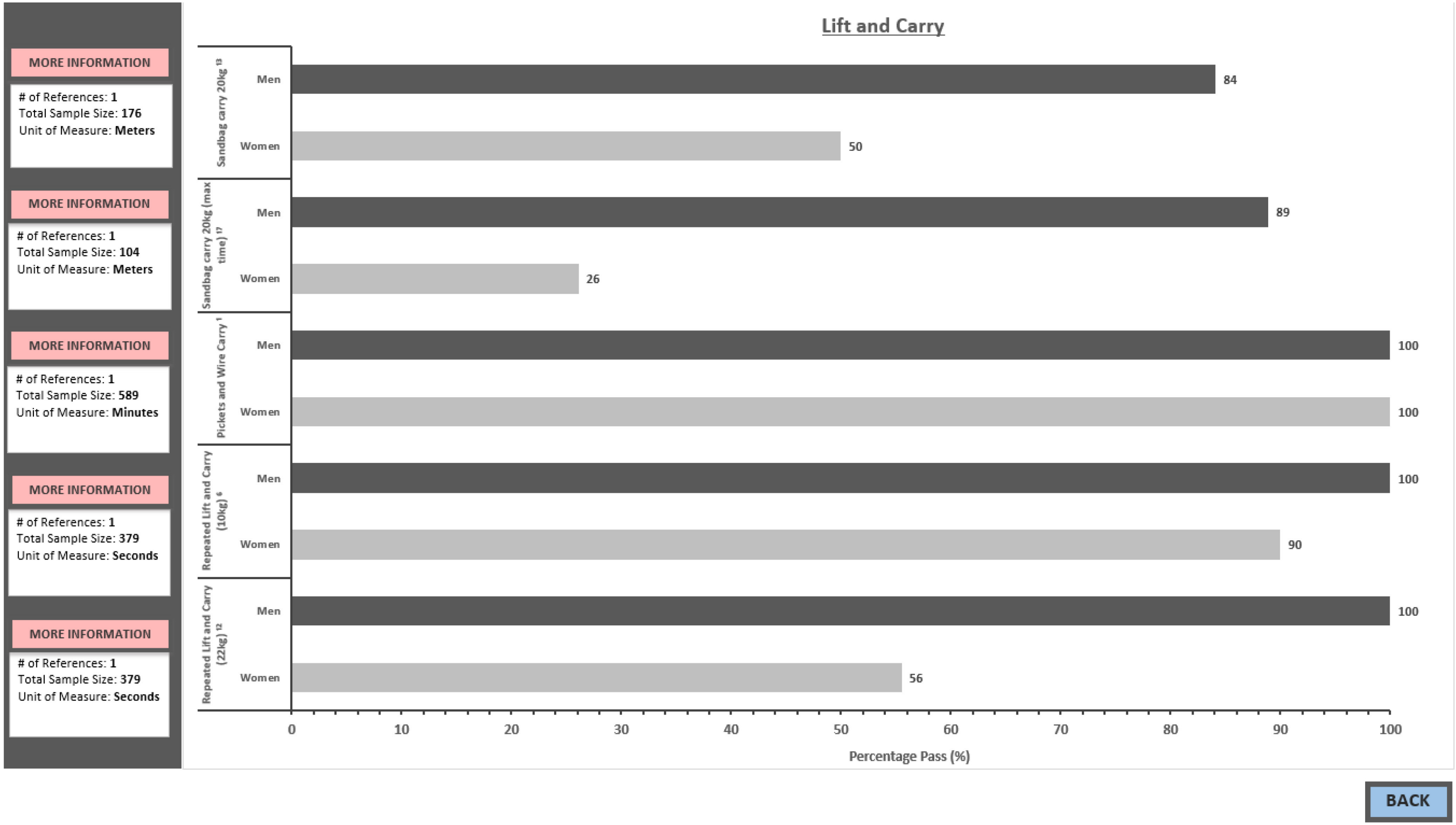
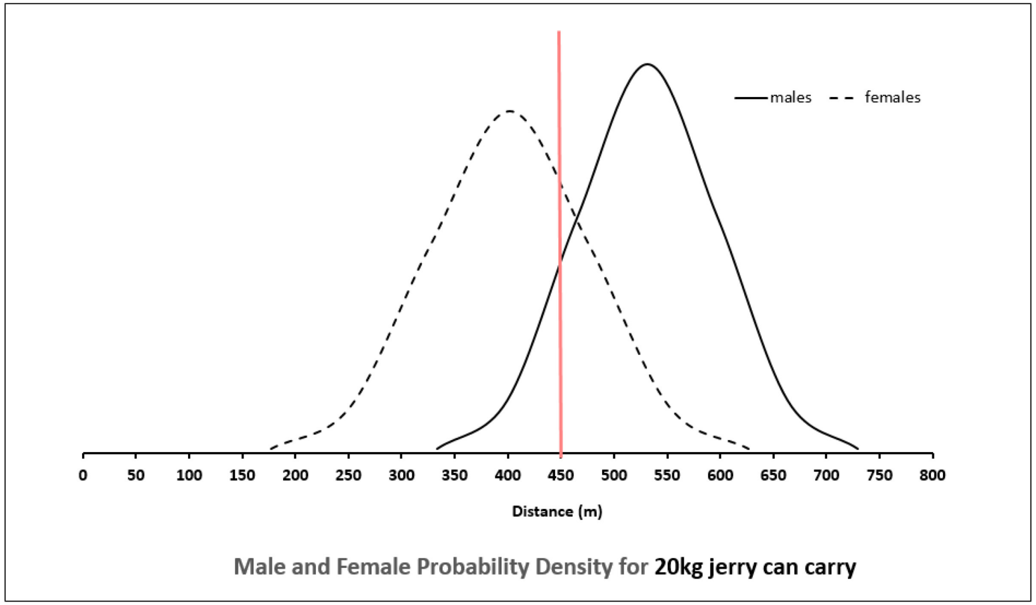


Figure 5-13: An Example of the Lift and Carry Breakdown, Comparing

Performance Between Men and Women in Different Studies.



Task Description
Participants carried a 20 kg sandbag no more than chest height for 30 m, then placed it on a 1.1 m platform as many times as possible in 10 min. There was a 30 m walk, jog or run recovery before picking up the next sandbag from the floor and repeating the 30 m carry task.

BACK

Reference	Country	Branch	Sample Size (TOTAL)	Males	Females	Predictor Tests	Cut Score	Units	Male Score (mean)	Male Score (StDev)	Female Score (mean)	Female Score (StDev)	M-F Overlap
Richmond et al 2008	UK	Royal Air Force	104	70	34	Sandbag carry	450	m	531	66	402	75	36.01%

Figure 5-14: An Example of the Density Curves of Male and Female Performance on the 20 kg Jerry Can Carry Task. Also given is a breakdown of the study the density curves are based on, as well as a description of the selection test for this task.

5.8 REFERENCES

[1] Israeli Defense Force: Caracal Battalion. Available at https://www.idf.il/en/minisites/caracal-battalion/.

[2] Greeves, J.P. (2015). Physiological Implications, Performance Assessment and Risk Mitigation Stratetgies of Women in Combat-Centric Occupations. Journal of Strength and Conditioning Research; 29: S94-S100, 2015.

[3] Epstein, Y., Yanovich, R., Moran, D.S., and Heled, Y. (2013). Physiological Employment Standards IV: Integration of Women in Combat Units Physiological and Medical Considerations. European Journal of Applied Physiology; 113: 2673-2690.

[4] Smith D.L., Fehling P.C., and Segrave J.O. (2008). Ch. 11: Exercise, Work and Stress in Adolescent and Adult Women. In Physiological Bases of Human Performance During Work and Exercise. Ed. N.A.S. Taylor and H. Groeller. Elsevier, London.

[5] Roberts, D., Gebhardt, D.L., Gaskill, S.E., Roy, T.C., and Sharp, M.A. (2016). Current Considerations Related to Physiological Differences Between the Sexes and Physical Employment Standards. Applied Physiology Nutrition and Metabolism; 41: S108-120.

[6] Reilly, T.J., Sharp, M.A., Cao, M., and Canino, M.C. (2019). A Database of Predictor Test Sex Bias for Development of Military Physical Employment Standards. Work; In Press.

[7] Gordon, C.C., Churchill, T., Clauser, C.E., Bradtmiller, B., McConville, J.T., Tebbetts, I., and Walker, R.A. (1988). Anthropometric Survey of U.S. Army Personnel: Methods and Summary Statistics. Technical Report Natick TR-89/027, U.S. Army Natick Soldier Research, Development and Engineering Center, Natick, MA, US.

[8] Malina, R.M. and Bouchard C. (1991). Growth, Maturation and Physical Activity, Human Kinetics Academics, Champaign, IL.

[9] Miller, A.E.J., MacDougall, J.D., Tarnopolsky, M.A., and Sale, D.G. (1993). Gender Differences in Strength and Muscle Fiber Characteristics. European Journal of Applied Physiology; 66(3): 254-262.

[10] Jackson, A.S. (1994). Preemployment Physical Evaluation. Exercise and Sport Science Reviews; 22: 53-90.

[11] Rayson, M. (1998). The Development of Physical Selection Procedures. Phase 1: Job Analysis. Presented at Annual Conference of the Ergonomics Society, Cirencester, UK.

[12] Sharp, M.A., Patton, J.F., and Vogel, J.A. (1998). A Database of Physically Demanding Tasks Performed by U.S. Army Soldiers. Technical Report T98-12, U.S. Army Research Institute of Environmental Medicine, Natick, MA.

[13] Wilmore, J.H., Costill, D.L., and Kenney, W.L. (2008). Physiology of Sport and Exercise, 4th ed., Human Kinetics, Champaign, IL.

[14] Nindl, B.C., Eagle, S.R., Frykman, P.N., Palmer, C., Lammi, E., Reynolds, K., Allison, K., and Harman, E.(2017). Functional physical Training Improves Women's Military Occupational Performance. Journal of Science and Medicine in Sport; 20: S91-S97.

[15] Sharp, M.A. (1994). Physical Fitness and Occupational Performance of Women in the US Army. Work; 4: 80-92.

[16] Reilly T.J., Stockbrugger B., Larocque S., Walsh E., and Gagnon P. (2017). A Fitness Assessment for the Canadian Army – FORCE Combat. Journal of Science and Medicine in Sport; 20: S127.

[17] Nindl, B.C., Mahar, M.T., Harman, E.A., and Patton, J.F (1995). Lower and Upper Body Anaerobic Performance in Male and Female Adolescent Athletes. Medicine and Science in Sports and Exercise; 27(1): 235-241.

[18] Yanovich, R., Evans, R., Israeli, E., Constantini, N., Sharvit, N., Merkel, D., Epstein, Y., and Moran, D.S. (2008). Differences in Physical Fitness of Male and Female Recruits in Gender-Integrated Army Basic Training. Medicine and Science in Sports and Exercise; 40: S654-S659.

[19] Vogel, J.A., Patton, J.F., Mello, R.P., and Daniels, W.L. (1986). An Analysis of Aerobic Capacity in a Large United States Population. Journal of Applied Physiology; 60(2): 494-500.

[20] Knapik J.J., Vogel J.A., and Wright J.E. (1981). Measurement of Isometric Strength in an Upright Pull of 38 cm. Technical Report #T3/81, U.S. Army Research Institute of Environmental Medicine Technical Report, Natick, MA.

[21] Foulis, S.A., Redmond,, J.E., Warr, B.J., Sauers, S., Walker, L., Canino, M., Hydren, J., Zambraski, E.J., Frykman, P.N., and Sharp, M.A. (2015). Development of a Physical Employment Testing Battery for Field Artillery Soldiers: 13B Cannon Crewman and 13F Fire Support Specialist. Technical Report T16-9, U.S. Army Research Institute of Environmental Medicine, Natick, MA, DTIC: AD10007859.

[22] Hodgdon, J.A. and Jackson A.S. (2000). Ch. 5: Physical Test Validation for Job Selection. In The Process of Physical Fitness Standards Development. Ed. S. Constable and B. Palmer. HSIAC Program Office, Wright-Patterson AFB, OH.

[23] Courtright, S.H., Mccormick, B.W., Postlethwaite, B.E., Reeves, C.J., and Mount, M.K. (2013). A Meta-Analysis of Sex Differences in Physical Ability: Revised Estimates and Strategies for Reducing Differences in Selection Contexts. Journal of Applied Psychology; 98: 623-641.

[24] Gebhardt, D.L. and Baker, T.A. (2010). Ch. 7: Physical Performance. In Handbook of Workplace Assessment. Ed. J.C. Scott and D.H. Reynolds. Jossey-Bass, San Francisco, CA, US.

[25] Gebhardt, D.L. (2000). Ch. 6: Establishing Performance Standards. In The Process of Physical Fitness Standards Development. Ed. S. Constable and B. Palmer. HSIAC Program Office, Wright-Patterson AFB, OH, US.

[26] Gebhardt, D.L. and Baker, T.A. (2010). Ch. 13: Physical Performance Tests. In Handbook of Employee Selection. Ed. J.L. Farr and N.T. Tippins. Routledge, N.Y., N.Y., US.

[27] Nindl, B.C. (2015). Physical Training Strategies for Military Women's Performance Optimization in Combat-Centric Occupations. The Journal of Strength & Conditioning Research, 29, S101-S106.

[28] Harman, E.A., Gutekunst, D.J., Frykman, P.N., Nindl, B.C., Alemany, J.A., Mello, R.P., and Sharp, M.A. (2008). Effects of Two Different Eight-Week Training Programs on Military Physical Performance. Journal of Strength and Conditioning Research; 22: 524-534.

[29] Knapik, J.J., Banderet, L.E., Vogel, J.A., Bahrke, M.S., and O'Connor, J.S. (1996). Influence of Age and Physical Training on Measures of Cardiorespiratory and Muscle Endurance. European Journal of Applied Physiology and Occupational Physiology; 72: 490-495.

[30] Kraemer, W.J. (2004). Effects of Concurrent Resistance and Aerobic Training on Load-Bearing Performance and the Army Physical Fitness Test. Military Medicine; 169: 994-999.

[31] Hendrickson, N.R., Sharp, M.A., Alemany, J.A., Walker, L.A., Harman, E.A., Spiering, B.A., Hatfield, D.L., Yamamoto, L.M., Maresh, C.M., Kraemer, W.J., and Nindl, B.C. (2010). Combined Resistance and Endurance Training Improves Physical Capacity and Performance on tactical Occupational Tasks. European Journal of Applied Physiology; 109: 1197-1208.

[32] Knapik, J.J. and Gerber J. (1996). The Influence of Physical Fitness Training on the Manual Material-Handling Capability and Road-Marching Performance of Female Soldiers. Aberdeen Proving Ground, MD, US.

[33] Kraemer, W.J., Mazzetti, S.A., Nindl, B.C., Gotshalk, L.A., Volek, J.S., Bush, J.A., Marx, J.O., Dohi, K., Gomez, A.L., Miles, M., Fleck, S.J., Newton, R.U., and Hakkinen, K. (2001). Effect of Resistance Training on Women's Strength/Power and Occupational Performances. Medicine and Science in Sports and Exercise; 33: 1011-1025.

[34] Santtila, M., Pihlainen, K., Viskari, J., and Kyrolainen, H. (2015). Optimal Physical Training During Military Basic Training Period. Journal of Science and Medicine in Sports; 29 Suppl 11: S154-157.

[35] Drain, J.R., Burley, S.D., and Groeller, H. (2016). Investigation of an Alternate Physical Conditioning Regimen for Australian Army Recruits. Fisherman's Bend, Victoria, AU.

[36] Knapik J.J. and Sharp M.A. (1998). Task-Specific and Generalized Physical Training for Improving Manual-Material Handling Capability. International Journal of Industrial Ergonomics; 22: 149-160.

[37] Jones, B.H., Hauret, K.G., Dye, S.K., Hauschild, V.D., Rossi, S.P., Richardson, M.D., and Friedl, K.E. (2017). Impact of Physical Fitness and Body Composition on Injury Risk Among Active Young Adults: A Study of Army Trainees. Journal of Science and Medicine in Sport; 20: S17-S22.

[38] Gumieniak, R.J., Gledhill, N., and Jamnik, V.K. (2018). Physical Employment Standard for Canadian Wildland Firefighters: Examining Test-Retest Reliability and the Impact of Familiarisation and Physical Fitness Training. Ergonomics; 61: 1324-1333.

[39] Teves, M.A., Wright, J.E., and Vogel, J.A. (1985): Performance on Selected Candidate Screening Test Procedures Before and After Army Basic and Advanced Individual Training. Natick, MA, US.

[40] Duran, A., Ozturk, H., Tekelioglu, U.Y., Ocak, T., and Emet, M. (2013). The Effect of Stretcher Type on Safety and Ease of Treatment in an Emergency Department. Turkish Journal of Medical Science; 43: 4.

[41] Reilly, T., Blacklock, R., Newton, P., Olinek, S., O'Hearn, K., and Spivock, M. (2013). Project FORCE Phase II Report: Physical Demands of Common, Essential, Physically Demanding Tasks in the CF. Assistant Deputy Minister (Science and Technology), Ottawa, Canada.

[42] Blacker, S.D., Myers, S.D., Coakley, S., Walker, E., and Hale, B. (2016). Development of Interim Infantry British Army physical Selection Standards for (Recruits) (PSS(R)) for Men and Women. TIN 3.179, Defence Human Capability Science and Technology Centre.

[43] Blacker, S.D., Wilkinson, D.M., Brown, P.I., Nevill, A.M., and Rayson, M.P. (2011). Revalidation of Physical Selection Standards for Recruits [PSS(R)] for Junior, Infantry and Soldier Pre-Conditioning Courses. Optimal Performance Ltd, Bristol, UK.

[44] Blacker, S.D., Wilkinson, D.M., Brown, P.I., and Rayson, M.P. (2011). Revalidation of Physical Selection Standards for Recruits [PSS(R)] Undertaking the Common Military Syllabus (Standard Entry) [CMS(SE)]. Optimal Performance Ltd, Bristol, UK.

[45] Blacker, S.D., Wilkinson, D.M., and Rayson, M.P. (2012). Development of Physical Selection Standards for Officer Cadets [PSS(O)]. Optimal Performance Ltd, Bristol, UK.

[46] Beck, B., Carstairs, G.L., Billing, D.C., Caldwell, J.N., and Middleton, K.J. (2017). Modifiable Anthropometric Characteristics are Associated with Unilateral and Bilateral Carry Performance. Journal of Strength and Conditioning Research; 31: 489-494.

[47] Rayson, M., Holliman, D., and Belyavin, A. (2000). Development of Physical Selection Procedures for the British Army. Phase 2: Relationship Between Physical Performance Tests and Criterion Tasks. Ergonomics; 43: 73-105.

[48] Hackfort, D., Leyk, D., and Scherer, H.G. (2017). Psychophysical Performance and Military Fitness in the Light of Mission-Requirements and Percent Level of Performance Levels of Female and Male Soldiers. Bonn, Germany.

[49] Rohde, U., Sievert, A., Rüther, T., Witzki, A., and Leyk, D. (2015). Concept for a Predeployment Assessment of Basic Military Fitness in the German Armed Forces. Journal of Science and Medicine in Sports; 29: S211-S215.

[50] Sharp, M., Pietila, D., Alemany, J., Rarick, K., Staab, J., Nindl, B., Kraemer, W., Spiering, B., Hatfield, D., and Maresh, C. (2006). Comparison of Three Training Programs for Improving Repetitive Lifting Task Performance in Women. Medicine and Science in Sports and Exercise; 38.

Chapter 6 – THE ROLE OF PHYSICAL EMPLOYMENT STANDARDS IN MUSCULOSKELETAL INJURY PREVENTION

K.G. Hauret
US Army Public Health Center
UNITED STATES

J. Drain
Defence Science and Technology Group
AUSTRALIA

A. Fieldhouse
UK Ministry of Defence
UNITED KINGDOM

T.J. Reilly
Department of National Defence
CANADA

S. Jackson
UK Ministry of Defence
UNITED KINGDOM

6.1 BACKGROUND

Historically, women have served in various roles during military conflicts including combat while disguised as men, camp domestics, and attending to the sick and wounded [1]. During the last half of the twentieth century, militaries across NATO began to actively integrate women into military service, but occupations were generally limited to non-combat occupations. However, in recent years many militaries have opened previously closed combat occupations to capable women [2], [3]. Personnel serving in ground combat occupations (e.g., infantry, armor, artillery, combat engineers) and combat support (e.g., military police, signal) occupations carry the heaviest loads for longer durations, and lift or carry the heaviest equipment [4], [5], [6], [7]. Military members in these occupations must be ready and able to perform these physically demanding tasks whenever called upon. These physically demanding tasks are challenging for all personnel, but the physical challenge is generally greater for women due to physical and physiological differences, particularly in muscular strength and power [3], [8], [9]. These activities are also associated with high Musculoskeletal Injury (MSKI) risks and these risks are generally higher among women [10], [11], [12], [13], [14], [15].

Across NATO, militaries are developing and implementing PES as a means to select and train individuals that are capable of safely and effectively performing essential physically demanding tasks, and be more resilient to injury. An anticipated secondary outcome of implementing PES is a reduction in the MSKI rates as military members develop the requisite physical capacity to meet the higher physical standards [3], [16], [17]. The purpose of this chapter is to describe:

1) The training-related MSKI problem for the military;
2) Risk factors for training-related injuries; and
3) The potential role of PES to influence changes in physical training regimes so that personnel can achieve the higher physical standards while also reducing their risk of injury.

6.2 EPIDEMIOLOGY OF MUSCULOSKELETAL INJURIES IN THE MILITARY

MSKIs are the biggest health problem confronting military forces in peacetime and combat operations [18]. As an example, across the military services in the United States (US) there are more than 1.95 million medical visits annually for injuries [18], [19]. By comparison, the second leading cause of medical visits, behavioural health, accounts for approximately 750,000 visits. Injuries also affect more military members than any other diagnostic category.

For militaries to successfully reduce injury risks and lower the MSKI rates, a comprehensive public health approach to injury prevention is required [20]. This is, in fact, the same approach used by public health professionals to investigate and stop disease outbreaks or develop treatment protocols for emerging health issues. The five steps of this public health approach to injury prevention are presented in Figure 6-1 [18]. Each step serves to answer a key question, such as “Is there an injury problem, and how big is it?” or “What causes the problem?” To answer these first two questions, public health professionals conduct injury surveillance to monitor injury rates and trends, and identify injury-related activities and mechanisms of injury. Then, research and field investigations are conducted to identify and evaluate potential risk factors for injuries. The next sections of this chapter describe some important findings from these first two steps of the process.

Five Key Public Health Questions		Five Steps of the Public Health Approach
1. Is there a problem and how big is it?	→	1. Surveillance
2. What causes the problem?	→	2. Research & field investigations
3. What works to prevent the problem?	→	3. Intervention trials & systematic reviews
4. Who needs to know and do what?	→	4. Program and policy implementation
5. How effective is what we have done?	→	5. Public health evaluations & monitoring

Figure 6-1: The Public Health Approach to Injury Prevention [18].

6.2.1 Musculoskeletal Injuries in the Military

MSKI is the leading cause of medical downgrade (e.g., limited or restricted duty), training attrition, discharge and medical compensation among military personnel. MSKIs reduce the ‘deployable’ workforce which impacts operational capability [21], [22], [23], [24], [25], [26]. This resultant loss of capability and/or personnel also creates a training burden due to the requirement to continuously replace personnel and develop new capability. During recent military deployments and combat operations, MSKIs directly affect the readiness and capability of the deployed force. Non-battle injuries have caused more medical evacuations (34%) from the combat theatre than any other cause including combat injuries [27], [28], [29], [30]. MSKI can degrade health, lead to long-term physical impairment and affect career progression. Among non-deployed military members in the US, MSKIs result in more than 25 million days of limited or restricted duty, account for 47 to 57% of fatalities, and 22 to 63% of all disability when members leave the military [31]. MSKIs also represent a significant strain on the healthcare system and are associated with considerable medical compensations costs. For the US Army, the injury-related cost of medical care, rehabilitation, wages and sick days was estimated at $1.5 billion per annum in 2013 [23].

MSKI rates are highest for both sexes during the initial basic training of new recruits compared to other stages of military career [8], [15], [32], [33], [34], [35], [36]. Since men and women perform the same training throughout initial training courses, both sexes are exposed to the same injury risks and hazards. During this initial basic training, MSKI injury rates (per 100 trainees/month (100 person-months)) have ranged from 9.7 to 16.4 for men and from 16.0 to 33.3 for women. The MSKI rate for women in basic training has historically been two times higher than the rate for men [12], [36], [37], [38], [39]. Injury data from incumbent military populations is less abundant than basic training, especially for women due to their smaller numbers and historic differences in occupational fields to which women were assigned. Injury rates (per 100 person-years) for incumbent (i.e., experienced) men in combat occupational fields range from 58 to 144 per 100 person-years [25], [26]. Overall population injury rates for men and women in the US Army in

2016 were 133 and 179 per 100 person-years, respectively [38]. In 2016, similar to years 2011 to 2015, the MSKI rate for women was 1.3 times higher than the rate for men. Unlike initial basic training where men and women train together and have similar injury risk exposures, incumbent military members have different occupational injury risks depending on sex, rank, occupation, and type of assigned unit [7], [24], [25], [26], [40], [41], [42]. Additional studies of incumbent military members are needed to evaluate and compare MSKI rates between men and women that have similar occupational injury risks and exposures.

Overuse injuries are often the most prevalent injury type observed and are generally attributed to the arduous and repetitive nature of military training [17], [22], [33], [43], [44], [45]. The lower limb and back are the most common locations in both recruits/trainees [15], [39], [46], [47] and incumbents [25], [48]. Differences in MSKI rates and patterns (types, locations) between sexes have also been reported [10], [13], [39], [49], [50], [51], but not in all instances [33].

6.2.2 Risk Factors for Musculoskeletal Injury

A compendium of potential risk factors for MSKI have been evaluated in many recruit and incumbent military populations. But understanding the aetiology of MSKI risk can still be difficult as there are numerous confounders in understanding the association between physical capacity and occupational MSKI risk, including training status, genetics, sex, age, occupational exposure(s), and previous injuries.

Risk factors associated with MSKI can be broadly categorized as intrinsic and extrinsic risk factors [12], [14], [52]. Intrinsic factors are individual characteristics (e.g., age, stature, body mass, body composition, sex, physical capacity, injury history) and extrinsic factors are conditions and demands placed on personnel (e.g., environment, equipment, training demands) [14]. The intrinsic and extrinsic risk factors can be further sub-categorized as non-modifiable, meaning actions cannot be taken by the individual (intrinsic) or others (extrinsic) to change the individual's level of risk, or modifiable, meaning actions can be taken by the individual (intrinsic) or others (extrinsic) to change the individual's level of risk (Table 6-1).

Table 6-1: Categorization of Common Risk Factors for Musculoskeletal Injury.

Intrinsic Risk Factors		Extrinsic Risk Factors	
Non-Modifiable	**Modifiable**	**Non-Modifiable**	**Modifiable**
Age	Aerobic capacity	Mission-essential tasks	Distance run, walked, or marched
Sex	Body composition	Terrain	Time to complete a task
Stature	Muscular strength	Environment	Weight of load lifted or carried
Anatomy (pes planus, Q-angle, genu varum)	Athletic skills and ability		Training volume (frequency, duration, intensity)
Prior injury	Tobacco use		Footwear

6.2.2.1 Intrinsic Risk Factors for Musculoskeletal Injury

Considerable research has demonstrated associations between intrinsic risk factors and MSKI incidence in military personnel (Table 6-2) [8], [12], [17], [36], [37], [39], [47]. Of the modifiable intrinsic risk factors, physical capacity (e.g., cardiorespiratory fitness, muscular strength and muscular endurance) is the most plastic and modifiable [12], [39], [47], [53]. Furthermore, improving these modifiable MSKI risk factors (physical capacity) can help to moderate non-modifiable risk factors such as age and sex.

Table 6-2: Summary of Intrinsic Risk Factors for Musculoskeletal Injury in Military Personnel.

Risk Factor	Study (Reference)	Population	Results
Age	Knapik et al. 2001 [39]	U.S. Army recruits; n = 756 males and 474 females.	Male recruits aged 25 – 35 years had an injury risk ratio of 1.4 (95% CI 1.0 – 2.1, p = 0.07), compared to younger recruits. Females showed no association between age and MSKI.
	Heir and Eide 1997 [35]	Norwegian Army male conscripts; n = 480.	Risk ratio of 1.9 (95% CI 1.23 – 2.91, p < 0.01) for conscripts over the age of 22 years.
	Jones, Cowan et al. 1993 [54]	U.S. Army male infantry trainees; n = 303.	Relative risk 1.8 (95% CI 1.2 – 2.6, p ≤ 0.05) for trainees ≥ 24 years old compared to trainees 17 – 19 years old.
	Knapik et al. 2006 [16]	U.S. Army recruits; n = 1078 males and 731 females.	Men: Risk ratio of 1.36 (95% CI 1.0 – 1.8, p = 0.04) for men >24 years old compared to the youngest trainees (17 – 19 years). Women: Risk ratio of 1.31 (95% CI 1.03 – 1.61, p = 0.01) for trainees aged 20 – 24 years compared to the youngest trainees (17 – 19 years).
	Roy, Knapik et al. 2012 [50]	U.S. Army soldiers; n = 593 (n = 536 males, 57 females).	Risk ratio of 1.33 (1.01 – 1.72, p < 0.05) for soldiers 30 – 34 years old compared to soldiers < 25 years old.
Sex	Knapik et al. 2001 [39]	U.S. Army recruits; n = 756 males and 474 females.	Female recruits showed an injury ratio of 2.1 compared to male recruits.
	Roy, Knapik et al. 2012 [50]	U.S. Army soldiers; n = 593 (n = 536 males, 57 females).	Female incumbent soldiers had a risk ratio of 1.4 (95% CI 1.09 – 1.71, p < 0.05) compared to men.
	Harwood, Rayson, and Nevill 1999 [33]; Trank et al. 2001 [55]	British Army Officer Cadets; n = 106 (68 males, 38 females).	Injury incidence was similar between male and female Army officer cadets.
	Hauret, Steelman et al. 2018 [56]	U.S. Army recruits (men: n = 54,256; women: 18,469).	Numerous studies have observed a ~ two to five-fold higher incidence of MSKI in women compared to men, in various settings.

Risk Factor	Study (Reference)	Population	Results
Sex (cont'd)	Hauret, Steelman et al. 2018 [56]	U.S. Army soldier population.	Injury rate ratio for women compared to men was 1.3 for each year 2011 to 2016.
Body mass and BMI	Jones, Cowan et al. 1993 [54]	U.S. Army male infantry trainees; n = 303.	No association between percent body fat or BMI and injury risk in infantry recruits.
	Knapik et al. 2001 [39]	U.S. Army recruits; n = 756 males and 474 females.	No association of body mass or BMI and MSKI risk in either sex.
	Knapik et al. 2006 [16]	U.S. Army recruits; n = 1078 males and 731 females.	No association of BMI and injury risk in male or female recruits.
	Heir and Eide 1997 [35]	Norwegian Army male conscripts; n = 480.	No relationship between high BMI (23.6 – 33.9) (compared to all others) and injury (RR 1.3, 95% CI 0.85 – 2.01, p NS), higher injury risk for those with low BMI (17.4 – 21.4) (compared to all others) (RR 1.6, 95% CI 1.06 – 2.43, $p < 0.05$).
	Blacker et al. 2008 [8]	British Army recruits; n = 13,417 (11,937 males, 1,480 females).	Recruits with the lowest BMI (17 – 20) had the greatest likelihood of medical referral ($p < 0.05$).
Body composition	Jones, Bovee, et al. 1993 [12]	U.S. Army recruits, n = 124 males, 186 females.	Measured height, weight, body fat mass and fitness at the beginning of basic training and screened medical records at the end of training. Men in the highest and lowest quartiles of BMI had 2.8 and 2.3 times higher injury risk, respectively, than men in the "average" BMI group. This was not found in female recruits.
	Jones, Cowan, et al. 1993 [54]	U.S. Army male infantry trainees, n = 303.	No relationship between body fat % and injury risk.
	Blacker et al. 2008 [8]	British Army recruits, n = 13,417 (11,937 males, 1,480 females).	Recruits with the highest body fat % (21 – 45 %) had the greatest likelihood of medical referral (HR 1.65, 95% CI 1.31 – 2.07, $p < 0.001$).

Risk Factor	Study (Reference)	Population	Results
Cardiorespiratory fitness	Knapik et al. 2001 [39]	U.S. Army recruits, n = 756 males and 474 females.	Male recruits in the slowest quartile in the 3.2 km run had a time-loss injury risk ratio of 1.6 (95% CI 1.0 – 2.4 p = 0.04) compared to the fastest quartile. The slowest two quartiles of female recruits were at an increased risk of injury (RR 1.6 – 1.9, 95% CI 1.0 – 2.8, p < 0.05).
	Knapik et al. 2006 [16]	U.S. Army recruits, n = 1078 males and 731 females.	Male recruits in the slowest quartile in the 1.6 km run had a risk ratio of 1.43 (95% CI 1.0 – 1.96 p = 0.03) compared to the fastest quartile. Female recruits in the slowest quartile had a two-fold increased risk of injury (RR 2.01, 95% CI 1.52 – 2.66, p < 0.01).
	Harwood, Rayson, and Nevill 1999 [33]	British Army Officer Cadets, n = 106 (68 males, 38 females).	Males and females with slower 2.4 km run times were more likely to sustain an injury (p < 0.05).
	Jones, Cowan et al. 1993 [54]	U.S. Army male infantry trainees, n = 303.	Recruits in the fourth (slowest) quintile for the 3.2 km run had twice the injury risk compared to the fastest quintile (RR 2.1, 95% CI 1.1 – 4.2, p ≤ 0.05), but interestingly, the slowest quintile did not demonstrate a difference (RR 1.6, 95% CI 0.7 – 3.4, p > 0.05).
	Blacker et al. 2008 [8]	British Army recruits, n = 13,417 (11,937 males, 1,480 females).	The slowest quintile for the 2.4 km run had a 6-fold greater chance of medical referral (HR 6.25, 95% CI 4.66 – 8.39, p < 0.001). After adjusting for physical fitness, gender was not an independent risk factor for injury.
	Fallowfield et al. 2018 [37]	British Air Force recruits, n = 990 men, 203 women in Phase I training.	Crude relative risk of injury was 1.77, 95% CI1.49 – 2.10). When adjusted for education, height, and lifestyle measures, RR was 1.44, 95% CI, 1.1301.83. But adding aerobic fitness to the model, RR was 0.59, 95% CI 0.42 – 0.83. Physical fitness was the most important independent risk factor, not sex.
	Hall 2017 [57]	British Army male recruits, n = 3446.	The slowest 20% for the 2.4 km run had a two-fold higher injury incidence (p < 0.001).

Risk Factor	Study (Reference)	Population	Results
Cardiorespiratory fitness (cont'd)	Heir and Eide 1997 [35]	Norwegian Army male conscripts, n = 480.	No relationship between initial 3000 m run time and injury risk (p>0.05).
Muscular strength	Hoffman et al. 1999 [43]	Israeli male military recruits, n = 136	Recruits who were 1 standard deviation below the population mean for leg strength had an increased risk of lower extremity stress fracture (RR 4.7, 95% CI 1.7 – 13.6, p < 0.05).
Muscular endurance	Knapik et al. 2001 [39]	U.S. Army recruits, n = 756 males and 474 females	Male recruits in the two lower quartiles for push-ups had almost a two-fold higher risk of a time-loss injury compared to the highest quartile (RR 1.8, 95% CI 1.2 – 2.8 p < 0.01). Female recruits in the two lower quartiles for push-ups had a time-loss injury risk ratio of 1.6 compared to the highest quartile (95% CI 1.1 – 2.5 p = 0.02).
	Jones, Cowan et al. 1993 [54]	U.S. Army male infantry trainees, n = 303	Recruits in the second, fourth and fifth quintiles for initial push-ups had ≥ twice the injury incidence (RR 2.0 – 3.5, p ≤ 0.05) compared to recruits who did the most push-ups (first quintile).
	Knapik et al. 2006 [16]	U.S. Army recruits, n = 1078 males and 731 females	Male recruits in the two lower quartiles for push-ups had an increased risk of injury compared to the highest quartile (RR 1.63 – 1.93, 95% CI 1.1 – 2.7 p < 0.01). Female recruits in the lowest quartile for push-ups had an increased risk of injury compared to the highest quartile (RR 1.48, 95% CI 1.14 – 1.93 p < 0.01).
Previous injury	Montgomery et al. 1989 [58]	U.S. Army male trainees attending the 6 month SEAL assault training program, n = 505	Descriptive analysis: Of those who developed stress fracture (6.3%) 44% had had a prior history of shin pain.
	Robinson et al. 2016 [59]	British Army infantry recruits, n = 1,810	Men with an injury in the past 12 months had a 1.19 times higher risk of a training injury (HR 1.199 (95% CI 1.01 – 1.39, p = 0.03).

Risk Factor	Study (Reference)	Population	Results
Previous injury (cont'd)	Wilkinson et al. 2011 [60]	British Army infantry soldiers, n = 660	Soldiers completed a lifestyle questionnaire and were then followed for 12 months. Injuries were identified from electronic medical records. In multivariate Cox regression models, soldiers with a previous lower limb injury had 1.49 times injury risk (HR: 1.49, 95% CI 1.19 – 1.87, $p < 0.01$) and soldiers with a previous lower back injury had 1.30 times higher injury risk (HR: 1.30, 95% CI 1.03 – 1.63, $p = 0.03$).
	Kucera et al. 2016 [13]	Cadets at US military academies (Army, Air Force, and Navy); n = 9,811	Cadets completed a questionnaire at the beginning of training and were followed for 4 years. Cadets with prior history of lower extremity injury were at 1.74 times higher risk of injury (men: RR 1.74, 95% CI 1.55 – 1.94; women: RR1.74, 95% CI 1.52 – 1.99).

The strongest and most consistent association with MSKI has been cardiovascular fitness [12], [16], [17], [47], [61], [62], [63]. Studies from British Army basic training demonstrated that recruits with a 2.4 km run predicted VO_{2max} of 45.7 to 47.4 $mL \cdot kg^{-1} \cdot min^{-1}$ had double the risk of injury and discharge compared to recruits with a predicted VO_{2max} of 54.1 $mL \cdot kg^{-1} \cdot min^{-1}$ [57]. Similar findings were observed in US Army recruits where male recruits with a $VO_{2max} \leq 46.6$ $mL \cdot kg^{-1} \cdot min^{-1}$ were 2.2 times more likely to sustain an injury than those with a $VO_{2max} \geq 53.1$ $mL \cdot kg^{-1} \cdot min^{-1}$; and female recruits with a $VO_{2max} \leq 37.0$ $mL \cdot kg^{-1} \cdot min^{-1}$ were 2.8 times more likely to sustain an injury than those with a $VO_{2max} \geq 40.8$ $mL \cdot kg^{-1} \cdot min^{-1}$ [39].

As noted earlier, in recruit training populations where men and women train together and have similar injury risk exposures, the injury risk for women is approximately two times higher than the risk for men [12], [38], [39]. However, the entry-level physical capacity (e.g., cardiorespiratory fitness, muscular strength) of female recruits is lower than that of male recruits [14], [33], [35], [64]. Most reports of two times higher risk for women did not control for the entry-level physical capacity or physical characteristics (e.g., height, body composition) as co-variates. There is a growing body of evidence showing that injury risk for female recruits is similar to that of male recruits in multivariable analyses that controlled for age, aerobic capacity, and other physical characteristics [4], [36], [37], [40], [65] [66]. In these models, aerobic capacity is the primary independent risk factor for MSKI, rather than sex. In other words, these studies show that men and women with the same relative aerobic capacity have similar injury risks during recruit training.

Age-related declines in physical capacity tend to manifest in the 40s, but are influenced by numerous factors including genetics, lifestyle, training status and general health [67]. In healthy sedentary populations, cardiorespiratory endurance and muscular strength decline at a rate of approximately 10% per decade [68], [69], [70], [71], [72]. These age-related declines in physical capacity may make it increasingly difficult for individuals to meet occupational physical performance standards (e.g., PES), especially for those who are not required to participate in regular physical training. Furthermore, these normal age-related declines in physical capacities may impact women more than men, given that women on average required to work at higher relative intensities to complete occupational training and tasks [4], [73]. Physical training, however, can mitigate age-related declines in physical and physiological capacity.

Physical capacity, occupational exposure (an extrinsic risk factor) and MSKI risk are closely linked, and should be considered co-dependent factors in managing the physical performance and mitigating MSKI risk of military personnel. Physical training programs that focus on occupational demands and associated MSKI risk factors are, therefore, essential in developing and maintaining a physically capable and resilient workforce. Physical employment standards have an important role in establishing physical standards for military members and informing the development of the physical training regimen for military personnel.

6.2.2.2 Extrinsic Risk Factors for Musculoskeletal Injury

Basic military training is typically biased towards prolonged, moderate-intensity exercise, and includes activities undertaken within physical training sessions (e.g., prolonged running, circuit training, obstacle courses and marching), as well as various other tactical and combat-related training activities such as navigation, section attacks and assault courses. The cumulative physical activity exposure from these training activities is typically high [7], [74], [75]. This high level of physical activity during military training has been repeatedly associated with MSKIs, and in particular overuse injuries [7], [74], [75]. In fact, two systematic reviews examining MSKI prevention in the military have recommended a reduction in physical activity volume as a priority strategy [76], [77]. Moreover, it may be suggested that the typical volume of running and load carriage during basic military training is a particularly strong moderator of MSKI rates in recruits [47], [54]. For example, Jones, et al. [54], showed that US Army recruits covering the greatest distance on foot during military training (i.e., marching and running) had the highest incidence of MSKI [54]. Similarly, US Navy recruits with the highest running mileage during basic training also demonstrated the highest injury rate, when compared to the groups with lower running mileage over the 7-week training period (22.4% vs 17.2%) [55]. Additionally, the two lowest total run mileage quartiles (11.5 to 17.5, and 18.0 to 21.5 total miles) demonstrated a faster 2.4. km run time at the end of training, when compared to the highest total run mileage quartile (25.5 to 43.5 total miles). Therefore, a lowering of cardiorespiratory (running) training volume may mitigate injury risk. More importantly, this lowered training volume does not have to compromise performance gains.

A reduction in running distance/duration and a concomitant increase in intensity (e.g., high-intensity intervals) has been shown to be effective in improving cardiorespiratory endurance in military personnel, despite significant reductions in training volume [78], [79]. For example, Westcott et al. [79] compared two different physical training regimens in US Air Force personnel who had failed the annual physical fitness assessment [79]. The control group undertook the standard Air Force 12-week remedial physical training program that involved mostly running for approximately 60 min, 4 – 5 days a week. The intervention group undertook a whole-body circuit training program that involved 60 sec bouts of moderate load strength training (40 – 60 % 1RM) alternating with 60 sec efforts on a cycle ergometer for 25 min, 3 days per week. The aerobic training group (i.e., standard 12-week remedial physical training program) showed no change in 2.4 km run time or 1-minute push-ups, whereas the circuit training group showed significant performance gains ($p < 0.05$).

Furthermore, there is evidence to support lower volume, higher intensity endurance exercise combined with resistance exercise for the simultaneous development of muscular strength and cardiorespiratory endurance [54], [80], [81], [82]. Strength training will not only help to condition personnel for the many manual material handling tasks they are likely to perform, but also enhance load carriage and cardiorespiratory endurance performance [83], [84], [85]. The Australian Army recently trialled a higher intensity, lower volume physical training program in recruits undertaking the 12-week basic training course. The experimental training program introduced dedicated strength training sessions, reduced the volume of cardiorespiratory endurance sessions (from 17 to 8 sessions), and replaced steady-state running with high-intensity intervals. Load carriage sessions within the physical training program were also reduced from 7 to 2. The ratio of strength training to endurance training was ~ 2:1. Importantly, the total number of physical training sessions and total training time were matched between the experimental program and the extant physical training program [86], [87]. Despite a ~50% reduction in both the number of endurance sessions and the physical activity volume within the sessions, the experimental program showed greater gains in cardiorespiratory endurance (12.9%) compared to the control group (8.1%). The experimental training regimen also achieved superior gains in muscular strength and occupationally-relevant tasks (load carriage, box lift), despite a reduced exposure to military circuits and load carriage sessions. Beyond performance gains, the injury incidence was reduced by ~40% in the experimental group compared to the control group. These results highlight the importance of developing underlying physical capacities, rather than simply replicating occupational tasks (e.g., load carriage) to both increase physical capacity and reduce MSKI risk [86], [87].

6.3 ROLE FOR PHYSICAL EMPLOYMENT STANDARDS IN MUSCULOSKELETAL INJURY PREVENTION

The process for military leaders to develop and implement PES was described in detail in Chapter 3. Once implemented, an evidenced based and defensible PES serves the following objectives:

- Provides a sex-neutral standard that all personnel must meet to enter the combat occupations.
- Modifies (lowers) the intrinsic MSKI risk of personnel by ensuring that personnel have the requisite physical capacity to perform the essential physically demanding tasks.
- Leads to a modification of the training regimen (extrinsic modifiable risk factor) as leaders and Subject Matter Experts (SME) apply evidence-based principles of physical training so military members of both sexes are able to meet the required physical performance standards and optimize physical capacity while simultaneously reducing the MSKI risk [3], [23], [88], [89], [90].

To demonstrate these objectives, a few examples of successfully implemented PESs are provided.

- **Australian Army**: The development of PES helped to characterize the acute and chronic exposures to occupational tasks (e.g., manual materials handling, load carriage) which informed both the physical training regimen and MSKI prevention programs. A recent PES review of 57 employment categories in the Australian Army identified 583 physically demanding tasks, of which 458 (~ 79%) were classified as manual materials handling tasks [88]. Load carriage was also a common requirement for military personnel. Evidence had demonstrated that these loaded activities (load carriage, manual material handling) pose an injury risk to personnel, from both acute and chronic exposure[48], [91], [92]. Importantly, the focus of PES and physical training evolved beyond improving performance on select physical fitness tests (e.g., push-ups, 2.4 km run) and discrete occupational tasks, towards improving physical and physiological resilience to acute and chronic occupational task demands.
- **Canadian Armed Forces.** Emerging data indicates that PES can contribute to a reduction in MSKI by better matching the demands of the job with the physical capacity of personnel. The Canadian Forces collected data on recruits since 2014 (n = 8609; 7265 males, 1344 females). Results showed that of those recruits who passed the PES, the bottom 10% of performers on a given element were three times more likely to sustain a MSKI during basic training [93]. In addition, those that failed the sandbag drag (a strength test) had a 6-times higher injury risk than those who passed [66]. In multivariable regression model, performance on the PES was an independent predictor of MSKI, but similar to results presented earlier, sex was not an independent risk factor for MSKI.
- **US Army.** In January 2017, the Army implemented the four-event Occupational Physical Assessment Test (OPAT) that included the interval aerobic run, seated power throw, standing long jump, and standing deadlift [94]. Three performance standards were established for the OPAT (Heavy, Significant, and Moderate) which matched the physical demand categories that had been assigned to all occupational specialties (Heavy, Significant, and Moderate). The combat occupations were categorized as "heavy" physical demands. Recruits enlisting in the combat occupational fields were required to meet the highest (Heavy) standard on the OPAT. MSKI risks were assessed during the recruits first ten weeks of basic training [56]. Male (n = 13,067) and female (n = 3,857) recruits that met the lowest performance standard (Moderate) on the four-event OPAT had a 1.23 ($p < 0.01$) and 1.07 ($p = 0.05$) times higher risk of MSKI, respectively, compared to within sex recruits that met the highest standard. Low performance was also associated with MSKI risk on each individual event for both sexes. By setting minimum physical standards for all recruits and by matching the physical capability of recruits to the occupational field, a reduction in MSKI rates is expected in the long-term.

6.4 FUTURE DIRECTIONS

6.4.1 Reassessment of Physical Employment Standards

As shown in Figure 6-1, the last two steps of the public health approach to injury prevention are implementation of programs and policies, followed by monitoring and re-evaluation to assess the effectiveness of interventions. Many NATO militaries implemented a PES after completing steps 1 through 3 of the public health process. Once implemented, scientists, clinicians, and military leaders must combine efforts to continuously monitor and re-evaluate the effectiveness of the PES. Measures of effectiveness include improved performance on the PES, and, more importantly, improved performance on physically demanding occupational tasks and lower MSKI rates. Based on the results of the PES reassessment, scientists and leaders may decide to change performance standards on the PES, or may decide to change some of the physical assessments in the PES. No matter the outcome of this reassessment, scientists and military leaders should continue to monitor and revaluate their PESs.

Future studies evaluating the effects of physical training programs to improve performance on the PES should measure the relative strain (percentage of maximal capacity) of the participant while performing PES tasks, in addition to the performance outcome (e.g., completion time for a loaded march). This data will provide more insight into an individual's ability to cope with occupational rigours. PESs typically reflect a minimum dichotomous pass/fail occupational requirement, however they can be used to assess performance improvements and identify MSKI risks when designed and administered to collect maximal effort as a continuous variable. With this evolution there may be more significant reductions in work-related MSKIs. Military organizations should also support participation in physical training throughout the service members' career to not only preserve occupationally-relevant physical performance and mitigate injury risk, but also maintain underlying health.

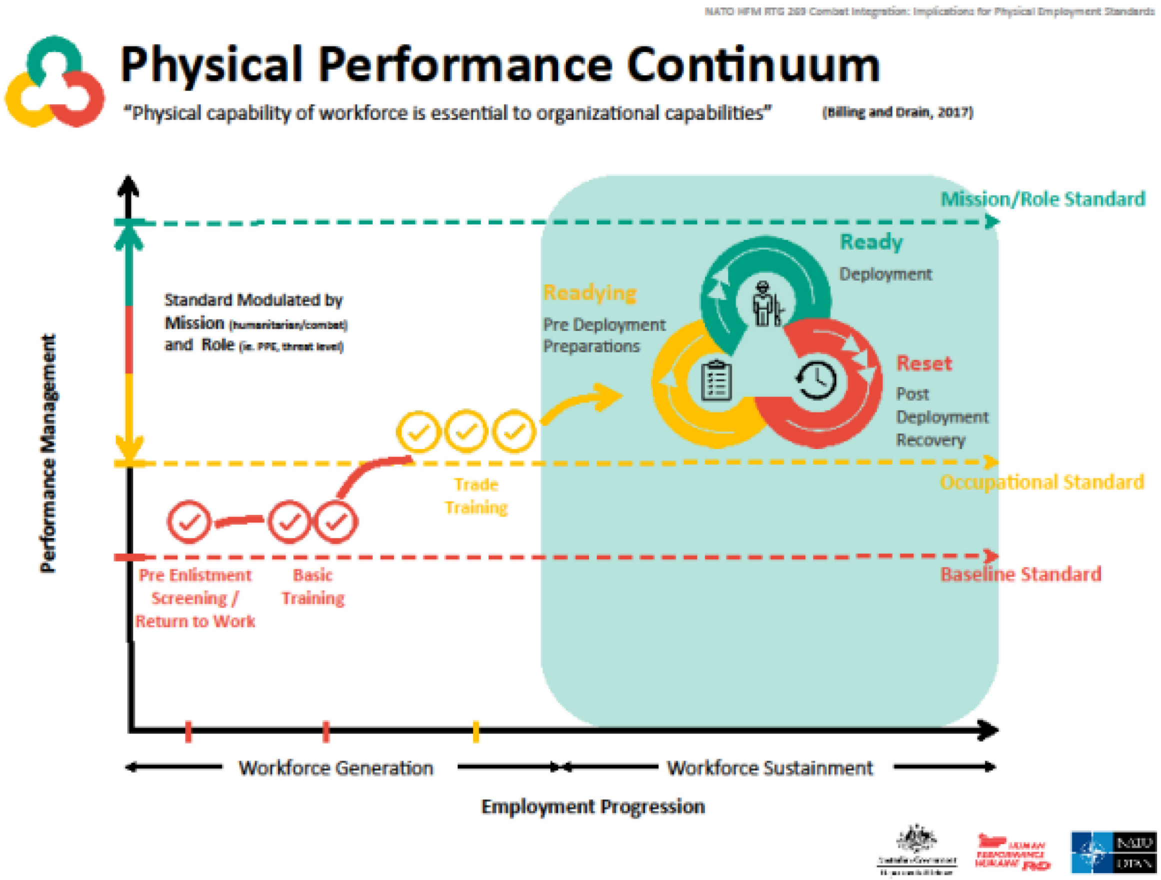


Figure 6-2:Through-Career Physical Performance Continuum (Adapted from Ref. [95]).

6.4.2 Strategic Plan to Prevent Musculoskeletal Injuries

MSKIs are not an homogenous group of injuries; therefore ascertaining the direct relationship of MSKIs with any particular mitigation is difficult. Scientists, clinicians, and military leaders should collaborate to develop and resource a broad strategic plan and programs to further reduce the impact of training-related MSKIs on the military force. Scientists must help their organization monitor injury rates and trends in the field force, and identify the leading activities and mechanisms of injury. They can then develop focused evidence-based strategies to mitigate the modifiable injury risk factors and target prevention efforts at leading causes of injury, and those injuries with greatest potential for morbidity, time-loss, and cost [20]. This injury prevention plan should be aligned with the five steps of the public health approach to injury prevention (Figure 6-1) [19].

6.4.3 Surveillance of Musculoskeletal Injuries and Physical Fitness

As the first step in the public health approach to injury prevention, scientists will conduct injury surveillance to monitor trends in MSKI rates, identify the most frequent and serious types of injuries, and identify the leading causes of injury (i.e., activities and mechanisms of injury). Surveillance may also include monitoring body composition (i.e., weight, BMI) and performance on physical assessments such as the PES. Scientists may conduct epidemiologic analyses of available surveillance data (e.g., injury data, demographics, body composition, physical assessments) to evaluate potential risk factors for injury. Table 6-3 summarizes the injury-related surveillance that is currently being conducted by some of the NATO militaries.

Table 6-3: Surveillance of Musculoskeletal Injuries and Physical Fitness Measures by the UK, Australia, USA, and Canada.

Surveillance Type	UK	Australia	USA	Canada
Injury Surveillance	**Surveillance:** Electronic medical record with "READ" coding.	**Surveillance:** Workplace Health and Safety Database; Medical System Database.	**Surveillance:** Electronic medical record with ICD-10 codes; End-of-training surveys; Electronic profile for limited duty days and mechanism of injury.	**Surveillance:** Electronic medical record with ICD-10 codes; Recruit Health Questionnaire; Annual Period Health Assessment; Physical Therapy Clinic notes.
Physical Fitness Measures	Yes	Yes; attendance and fitness tests.	Yes, fitness tests and body composition.	Yes, maximal fitness tests and body composition.

Most militaries use electronic medical records in the delivery of primary and/or secondary health care to service personnel. Most rely on a version of the International Classification of Diseases (ICD) to code the diagnosis of injuries and diseases (WHO). Whilst these records and diagnostic codes are undoubtedly important sources of epidemiologic data, their main functions are to enable delivery of health care and provide a basis for billing. The ICD diagnostic codes assigned to any specific injury can vary widely based on the medical specialty of the care provider and experience of trained coders. This is especially true for overuse MSKIs. In most cases, the electronic medical records also lack detailed information about the causes of injury (i.e., injury-related activity and mechanism of injury).

To overcome these limitations of electronic medical records for injury surveillance, some militaries have developed alternative data sources for surveillance (Table 6-3). Some have established bespoke injury records (e.g., clinic notes by physical therapists and athletic trainers) that are proving to be more useful in terms of understanding changes in MSKIs and injury risks over time. As an example, Canadian Armed Forces provided electronic tablets to athletic trainees enabling them to record details about how injuries occurred close to the time of occurrence. Injury surveys are also being used to obtain valid, reliable information about injuries, especially when they are completed close to the time of injury. Surveys can be used to obtain:

1) Type of injury;
2) Whether the individual sought medical care;

3) Number of injury-related days of limited duty;
4) Activities and mechanism of injury;
5) Occupation;
6) Use of nutritional and performance-enhancing supplements; and
7) Smoking history and other risk factors [72], [96], [97].

Figure 6-3 is an example of a training-related injury survey being used by the Canadian Armed Forces [71], [97].

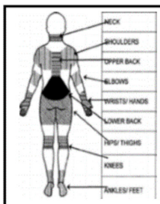
		Injured since your **last FORCE** evaluation	Did you have to be on modified duty? 0: no 1: Med-Chit, 2: Temporary Medical Employment Limitation 3: Permanent Medical Employment Limitation	Did you report your injury? 0: no 1: CoC only 2: Medical clinic 3: Physiotherapy	How did this injury limit activities? 0 : no limitations/ Injury 1: Minimally limited 2: Moderately limited 3: Significantly limited	What type of injury is it? 0: No injury 1: Overuse 2: traumatic 3: Exertion	What activity caused the injury? 0: Activities of daily living *Examples: Climbing stairs, tripping* 1: Military occupational demands *Examples: Rucksack march, casualty rescue, Camp construction* 2: Unit Sports/Recreational or Physical Training *Examples: Ball Hockey, soccer, morning Physical Training* 3: Physical training *Examples: Weight lifting, running, circuit training* 4: PSP organized sports *Examples: Triathlon, Ice Hockey, etc.*
1	Neck	Yes No	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3 4
2	Upper Back	Yes No	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3 4
3	Shoulders	Yes No	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3 4
4	Elbows	Yes No	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3 4
5	Wrists/ Hands	Yes No	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3 4
6	Lower back	Yes No	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3 4
7	Hips/ Thighs	Yes No	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3 4
8	Knees	Yes No	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3 4
9	Ankle / Feet	Yes No	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3	0 1 2 3 4

Figure 6-3: Sample Survey for Training-Related Musculoskeletal Injuries [71], [96].

6.4.4 Musculoskeletal Injury Research

Research is the 2nd step in the public health approach to injury prevention (Figure 6-1) [18]. Research is the means by which scientists generate a causal hypothesis for injury and then design a systematic inquiry to evaluate the hypothesis. Injury research can provide information on the incidence of injuries, injury causes, and risk factors.

A common study design in injury research is the prospective longitudinal study. With this study design, scientists follow a cohort of military personnel for a defined period of time to evaluate the association of risk factors and incident injuries. As an example, the Canadian Armed Forces is conducting a study in which military members will be followed for 5 years from recruitment to investigate the relationship between demographics, anthropometric measurements, operational fitness, and MSKIs. Figure 6-4 illustrates the injury data collection model for this Canadian study.

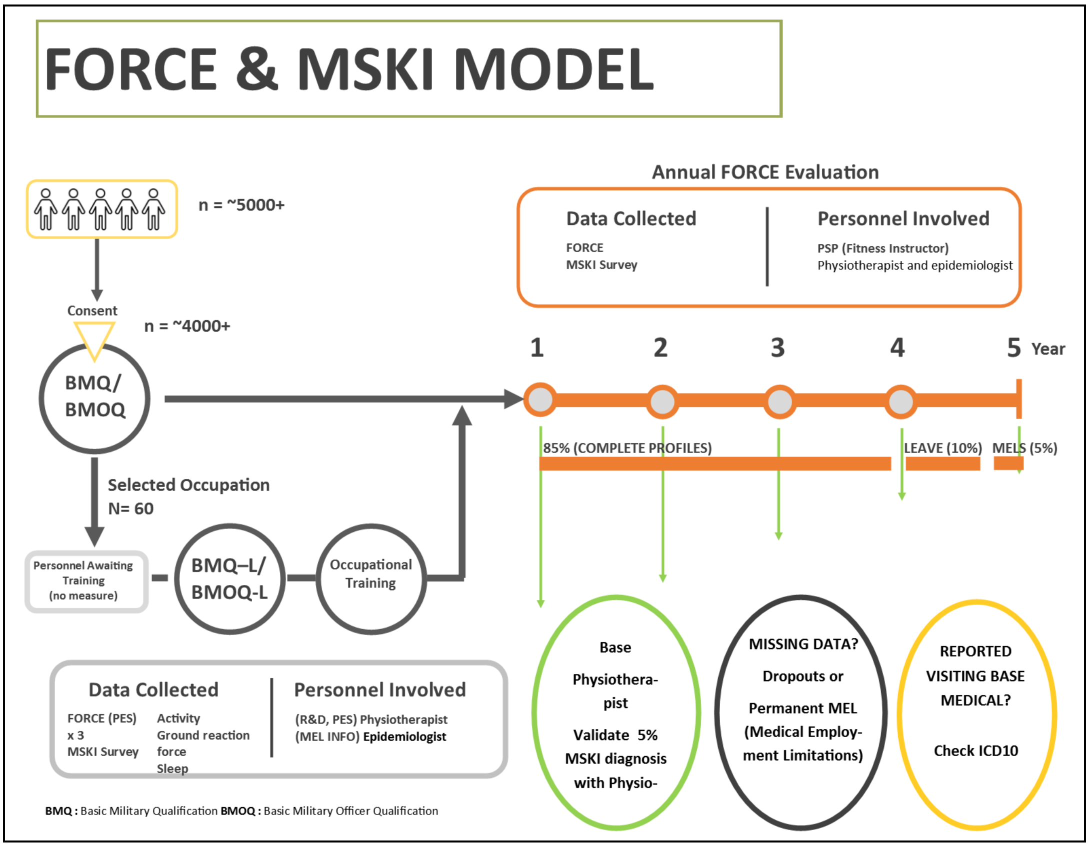


Figure 6-4: Injury Data Collection Model from Longitudinal Study by Canadian Armed Forces[66], [96].

6.5 CONCLUSION

The development of PES helps to characterize the acute and chronic exposure to occupational tasks (e.g., manual materials handling, load carriage) which can then help to calibrate injury prevention strategies and physical training programs. The focus of PES and physical training should evolve beyond improving performance on select fitness measures and discrete occupational tasks, towards improving physical and physiological resilience to acute and chronic occupational task demands. Future studies evaluating the effects of a training program on occupational task performance (or PES), should measure relative strain (percentage of maximal capacity) of the participant while performing the task, in addition to the performance outcome (e.g., completion time for a loaded march). This data will provide more insight into an individual's ability to cope with occupational rigours. PES typically reflect a minimum dichotomous pass/fail occupational requirement, however they can be used to assess performance improvements and identify MSKI risks when designed and administered to collect maximal effort as a continuous variable. With this evolution there may be more significant reductions in work-related MSKI. Furthermore, physically demanding occupations should support the through-career participation in physical training to not only preserve occupationally-relevant physical performance, but also cardiovascular health, while at the same time reducing MSKI risks. Physical training, PES and injury prevention strategies are co-dependent factors, and central to organizational capability within physically demanding occupations.

The introduction of a sex-neutral scientifically developed PES which reflects the physical demands of a job should result in a reduction in MSKI as a person-job fit is achieved at selection and maintained throughout a career. This will impact on what physical training service personnel undertake and should therefore be protective. As NATO militaries develop and implement PES, it is imperative that they develop an injury surveillance program, establish baseline injury rates, and then carefully monitor trends in injury rates, types, causes, and outcomes such as restricted duty after implementation of the PES.

6.6 REFERENCES

[1] Carreiras, H. (2006). Gender and the Military: Women in the Armed Forces of Western Democracies. Routledge, New York, NY, US.

[2] Greeves, J.P. (2015). Physiological Implications, Performance Assessment and Risk Mitigation Strategies of Women in Combat-Centric Occupations. J Strength Cond Res 29 Suppl 11:S94-100. doi: 10.1519/jsc.0000000000001116.

[3] Nindl, B.C., Jones, B.H., Van Arsdale, S.J., Kelly, K., and Kraemer, W.J. (2016). Operational Physical Performance and Fitness in Military Women: Physiological, Musculoskeletal Injury, and Optimized Physical Training Considerations for Successfully Integrating Women into Combat-Centric Military Occupations. Military Medicine, 181(suppl_1), 50-62.

[4] Blacker, S.D., Wilkinson, D.M., and Rayson, M.P. (2009). Gender Differences in the Physical Demands of British Army Recruit Training. Military Medicine, 174(8), 811-816.

[5] Foulis, S.A., Sharp, M.A., Redmond, J.E., Frykman, P.N., Warr, B.J., Gebhardt, D.L., ... and Zambraski, E.J. (2017). US Army Physical Demands study: Development of the Occupational Physical Assessment Test for Combat Arms Soldiers. Journal of Science and Medicine in Sport, 20, S74-S78.

[6] Sharp, M.A., Patton, J.F., and Vogel, J.A. (1996, October). A Data Base of Physically Demanding Tasks Performed by US Army Soldiers. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 40, No. 13, pp. 673-677). SAGE Publications, CA, US.

[7] Wilkinson, D.M., Rayson, M.P., and Bilzon, J.L. (2008). A Physical Demands Analysis of the 24-Week British Army Parachute Regiment Recruit Training Syllabus. Ergonomics, 51(5), 649-662.

[8] Blacker, S.D., Wilkinson, D.M., Bilzon, J.L., and Rayson, M.P. (2008). Risk Factors for Training Injuries Among British Army Recruits. Military Medicine, 173(3), 278-286.

[9] Epstein, Y., Fleischmann, C., Yanovich, R., and Heled, Y. (2015). Physiological and Medical Aspects that put Women Soldiers at Increased Risk for Overuse Injuries. The Journal of Strength & Conditioning Research, 29, S107-S110.

[10] Billings, C.E. (2004). Epidemiology of Injuries and Illnesses During the United States Air Force Academy 2002 Basic Cadet Training Program: Documenting the Need for Prevention. Military Medicine, 169(8), 664-670.

[11] Holsteen, K.K., Choi, Y.S., Bedno, S.A., Nelson, D.A., and Kurina, L.M. (2017). Gender Differences in Limited Duty Time for Lower Limb Injury. Occupational Medicine, 68(1), 18-25.

[12] Jones, B.H., Bovee, M.W., Harris III, J.M., and Cowan, D.N. (1993). Intrinsic Risk Factors for Exercise-Related Injuries Among Male and Female Army Trainees. The American Journal of Sports Medicine, 21(5), 705-710.

[13] Kucera, K.L., Marshall, S.W., Wolf, S.H., Padua, D.A., Cameron, K.L., and Beutler, A.I. (2016). Association of Injury History and Incident Injury in Cadet Basic Military Training. Medicine and Science in Sports and Exercise, 48(6), 1053.

[14] Neely, F.G. (1998). Intrinsic Risk Factors for Exercise-Related Lower Limb Injuries. Sports Medicine, 26(4), 253-263.

[15] Piantanida, N.A., Knapik, J.J., Brannen, S., and O'Connor, F. (2000). Injuries during Marine Corps Officer Basic Training. Military Medicine, 165(7), 515-520.

[16] Knapik, J.J., Darakjy, S., Hauret, K.G., Canada, S., Scott, S., Rieger, W., ... and Jones, B.H. (2006). Increasing the Physical Fitness of Low-Fit Recruits Before Basic Combat Training: An Evaluation of Fitness, Injuries, and Training Outcomes. Military Medicine, 171(1), 45-54.

[17] Rosendal, L., Langberg, H., Skov-Jensen, A., and Kjær, M. (2003). Incidence of Injury and Physical Performance Adaptations During Military Training. Clinical Journal of Sport Medicine, 13(3), 157-163.

[18] Jones, B.H., Canham-Chervak, M., and Sleet, D.A. (2010). An Evidence-Based Public Health Approach to Injury Priorities and Prevention: Recommendations for the US Military. American Journal of Preventive Medicine, 38(1), S1-S10.

[19] Jones, B.H., Canham-Chervak, M., Canada, S., Mitchener, T.A., and Moore, S. (2010). Medical Surveillance of Injuries in the US Military: Descriptive Epidemiology and Recommendations for Improvement. American Journal of Preventive Medicine, 38(1), S42-S60.

[20] Canham-Chervak, M., Hooper, T.I., Brennan Jr, F.H., Craig, S.C., Girasek, D.C., Schaefer, R.A., ... and Jones, B.H. (2010). A Systematic Process to Prioritize Prevention Activities: Sustaining Progress Toward the Reduction of Military Injuries. American Journal of Preventive Medicine, 38(1), S11-S18.

[21] Amoroso, P.J., Bell, N.S., and Jones, B.H. (1997). Injury Among Female and Male Army Parachutists. Aviation, Space, and Environmental Medicine, 68(11), 1006-1011.

[22] Finestone, A.S., Milgrom, C., Yanovich, R., Evans, R., Constantini, N., and Moran, D.S. (2014). Evaluation of the Performance of Females as Light Infantry Soldiers. BioMed Research International, 2014.

[23] Nindl, B.C., Williams, T.J., Deuster, P.A., Butler, N.L., and Jones, B.H. (2013). Strategies for Optimizing Military Physical Readiness and Preventing Musculoskeletal Injuries in the 21st Century. US Army Med Dep J:5-23.

[24] Reynolds, K., Cosio-Lima, L., Bovill, M., Tharion, W., Williams, J., and Hodges, T. (2009). A Comparison of Injuries, Limited-Duty days, and Injury Risk Factors in Infantry, Artillery, Construction Engineers, and Special Forces Soldiers. Military Medicine, 174(7), 702-708.

[25] Smith, T.A., and Cashman, T.M. (2002). The Incidence of Injury in Light Infantry Soldiers. Military Medicine, 167(2), 104-108.

[26] Teyhen, D.S., Goffar, S.L., Shaffer, S.W., Kiesel, K., Butler, R.J., Tedaldi, A.M., ... and Plisky, P.J. (2018). Incidence of Musculoskeletal Injury in US Army Unit Types: A Prospective Cohort Study. Journal of Orthopaedic & Sports Physical Therapy, 48(10), 749-757.

[27] Hauret, K.G., Pacha, L., Taylor, B.J., and Jones, B.H. (2016). Surveillance of Disease and Nonbattle Injuries During US Army Operations in Afghanistan and Iraq. US Army Med Dep J (2-16):15-23.

[28] Hauret, K.G., Taylor, B.J., Clemmons, N.S., Block, S.R., and Jones, B.H. (2010). Frequency and Causes of Nonbattle Injuries Air Evacuated from Operations Iraqi Freedom and Enduring Freedom, U.S. Army, 2001 – 2006. Am J Prev Med 38 (1 Suppl):S94-107. doi: 10.1016/j.amepre.2009.10.022. Available at: https://www.ncbi.nlm.nih.gov/pubmed/20117605.

[29] Skeehan, C.D., Tribble, D.R., Sanders, J.W., Putnam, S.D., Armstrong, A.W., and Riddle, M.S. (2009). Nonbattle Injury Among Deployed Troops: An Epidemiologic Study. Military Medicine, 174(12), 1256-1262.

[30] Waterman, B.R., Schoenfeld, A.J., Holland, C.A., Goodman, G.P., and Belmont Jr, P.J. (2011). Burden of Musculoskeletal Disease and Nonbattle Nontraumatic Injury in Both War and Disaster Zones. William Beaumont Army Medical Center El Paso TX Dept of Orthopaedic Surgery and Rehabilitation.

[31] Ruscio, B.A., Jones, B.H., Bullock, S.H., Burnham, B.R., Canham-Chervak, M., Rennix, C.P., ... and Smith, J.W. (2010). A Process to Identify Military injury Prevention Priorities Based on Injury Type and Limited Duty Days. American Journal of Preventive Medicine, 38(1), S19-S33.

[32] Almeida, S.A., Williams, K.M., Shaffer, R.A., and Brodine, S.K. (1999). Epidemiological Patterns of Musculoskeletal Injuries and Physical Training. Medicine and Science in Sports and Exercise, 31(8), 1176-1182.

[33] Harwood, G.E., Rayson, M.P., and Nevill, A.M. (1999). Fitness, Performance, and Risk of Injury in British Army Officer Cadets. Military Medicine, 164(6), 428-434.

[34] Heagerty, R., Sharma, J., Cayton, J., and Goodwin, N. (2018). Retrospective Analysis of Four-Year Injury Data from the Infantry Training Centre, Catterick. Journal of the Royal Army Medical Corps, 164(1), 35-40.

[35] Heir, T., and Eide, G. (1996). Age, Body Composition, Aerobic Fitness and Health Condition as Risk Factors for Musculoskeletal Injuries in Conscripts. Scandinavian Journal of Medicine & Science in Sports, 6(4), 222-227.

[36] Bell, N.S., Mangione, T.W., Hemenway, D., Amoroso, P.J., and Jones, B.H. (2000). High Injury Rates Among Female Army Trainees: A Function of Gender?. American Journal of Preventive Medicine, 18(3), 141-146.

[37] Fallowfield, J.L., Leiper, R.G., Shaw, A.M., Whittamore, D.R., Lanham-New, S.A., Allsopp, A.J., ... and Sanchez-Santos, M.T. (2018). Risk of Injury in Royal Air Force Training: Does Sex Really Matter? Military Medicine.

[38] Hauret, K.G., Canham-Chervak, M., Schuh-Renner, A., Taylor, B.J., Pierce, J.R., Jones, B.H., and Dada, E.O. (2018). Injury Surveillance and Longitudinal Studies for Gender Integration in the Army, Second Annual Assessment, 2017 (No. PHR-S. 0047783-17). US Army Public Health Center Aberdeen Proving Ground United States.

[39] Knapik, J.J., Sharp, M.A., Canham-Chervak, M., Hauret, K., Patton, J.F., and Jones, B.H. (2001). Risk Factors for Training-Related Injuries Among Men and Women in Basic Combat Training. Medicine and Science in Sports and Exercise, 33(6), 946-954.

[40] Anderson, M.K., Grier, T., Dada, E.O., Canham-Chervak, M., and Jones, B.H. (2017). The Role of Gender and Physical Performance on Injuries: An Army Study. American Journal of Preventive Medicine, 52(5), e131-e138.

[41] Knapik, J., Ang, P., Reynolds, K., and Jones, B. (1993). Physical Fitness, Age, and Injury Incidence in infantry Soldiers. Army Research Inst. of Environmental Medicine, Natick, MA, US.

[42] Roy, T.C., and Lopez, H.P. (2013). A Comparison of Deployed Occupational Tasks Performed by Different Types of Military Battalions and Resulting Low Back Pain. Military Medicine, 178(8), e937-e943.

[43] Hoffman, J.R., Chapnik, L., Shamis, A., Givon, U., and Davidson, B. (1999). The Effect of Leg Strength on the Incidence of Lower Extremity Overuse Injuries During Military Training. Military Medicine, 164(2), 153-156.

[44] Jordaan, G., and Schwellnus, M.P. (1994). The Incidence of Overuse Injuries in Military Recruits During Basic Military Training. Military Medicine, 159(6), 421-426.

[45] Ross, J. (1993). A Review of Lower Limb Overuse Injuries During Basic Military Training. Part 1: Types of Overuse Injuries. Military Medicine, 158(6), 410-415.

[46] Heir, T., and Glomsaker, P. (1996). Epidemiology of Musculoskeletal Injuries Among Norwegian Conscripts Undergoing Basic Military Training. Scandinavian Journal of Medicine & Science in Sports, 6(3), 186-191.

[47] Jones, B. H., Cowan, D. N., and Knapik, J. J. (1994). Exercise, Training and Injuries. Sports Medicine, 18(3), 202-214.

[48] Roy, T.C., Ritland, B.M., Knapik, J.J., and Sharp, M.A. (2012). Lifting Tasks are Associated with Injuries During the Early Portion of a Deployment to Afghanistan. Military Medicine, 177(6), 716-722.

[49] Armstrong I.D.W., Rue, J.P.H., Wilckens, J.H., and Frassica, F.J. (2004). Stress Fracture Injury in Young Military Men and Women. Bone, 35(3), 806-816.

[50] Roy, T.C., Knapik, J.J., Ritland, B.M., Murphy, N., and Sharp, M.A. (2012). Risk Factors for Musculoskeletal Injuries for Soldiers Deployed to Afghanistan. Aviation, Space, and Environmental Medicine, 83(11), 1060-1066.

[51] Wentz, L., Liu, P.Y., Haymes, E., and Ilich, J.Z. (2011). Females Have a Greater Incidence of Stress Fractures than Males in Both Military and Athletic Populations: A Systemic Review. Military Medicine, 176(4), 420-430.

[52] Mcleod, R., Stockwell, T., Rooney, R., Stevens, M., Phillips, M., and Jelinek, G. (2003). The Influence of Extrinsic and Intrinsic Risk Factors on the Probability of Sustaining an Injury. Accident Analysis & Prevention, 35(1), 71-80.

[53] Lisman, P.J., Sarah, J., Gribbin, T.C., Jaffin, D.P., Murphy, K., and Deuster, P.A. (2017). A Systematic Review of the Association Between Physical Fitness and Musculoskeletal Injury Risk: Part 1 – Cardiorespiratory Endurance. The Journal of Strength & Conditioning Research, 31(6), 1744-1757.

[54] Jones, B.H., Cowan, D.N., Tomlinson, J.P., Robinson, J.R., Polly, D.W., and Frykman, P.N. (1993). Epidemiology of Injuries Associated with Physical Training Among Young Men in the Army (No. USARIEM-M78-91). Army Research Inst of Environmental Medicine Natick MA.

[55] Trank, T.V., Ryman, D.H., Minagawa, R.Y., Trone, D.W., and Shaffer, R.A. (2001). Running Mileage, Movement Mileage, and Fitness in Male US Navy Recruits (No. NHRC-99-34). Naval Health Research Centre, San Diego, CA, US.

[56] Hauret, K. H., Steelman, R.A., Pierce, J.R. Alemany, J.A., Sharp, M.A., Foulis, S.A., & East, W.B. (2018). Association of Performance on the Occupational Physical Assessment Test (OPAT), Injuries, and Attrition during Initial Entry Training-OPAT Phase 1 (No. PHR No. S. 0047229-18b). Army Public Health Command Aberdeen Proving Ground MD Aberdeen Proving Ground United States.

[57] Hall, L.J. (2017). Relationship Between 1.5-Mile Run Time, Injury Risk and Training Outcome in British Army Recruits. Journal of the Royal Army Medical Corps, 163(6), 376-382.

[58] Montgomery, L.C., Nelson, F.R., Norton, J.P., and Deuster, P.A. (1989). Orthopedic History and Examination in the Etiology of Overuse Injuries. Medicine and Science in Sports and Exercise, 21(3), 237-243.

[59] Robinson, M., Siddall, A., Bilzon, J., Thompson, D., Greeves, J., Izard, R., and Stokes, K. (2016). Low Fitness, Low Body Mass and Prior Injury Predict Injury Risk During Military Recruit Training: A Prospective Cohort Study in the British Army. BMJ Open Sport & Exercise Medicine, 2(1), e000100.

[60] Wilkinson, D.M., Blacker, S.D., Richmond, V.L., Horner, F.E., Rayson, M.P., Spiess, A., and Knapik, J.J. (2011). Injuries and Injury Risk Factors Among British Army Infantry Soldiers During Predeployment Training. Injury Prevention, 17(6), 381-387.

[61] Pope, R.P., Herbert, R., Kirwan, J.D., and Graham, B.J. (1999). Predicting Attrition in Basic Military Training. Military Medicine, 164(10), 710-714.

[62] Taanila, H.P., Suni, J.H., Pihlajamäki, H.K., Mattila, V.M., Ohrankämmen, O., Vuorinen, P., and Parkkari, J.P. (2012). Predictors of Low Back Pain in Physically Active Conscripts with Special Emphasis on Muscular Fitness. The Spine Journal, 12(9), 737-748.

[63] Välimäki, V.V., Alfthan, H., Lehmuskallio, E., Löyttyniemi, E., Sahi, T., Suominen, H., and Välimäki, M.J. (2005). Risk Factors for Clinical Stress Fractures in Male Military Recruits: A Prospective Cohort Study. Bone, 37(2), 267-273.

[64] Heller, R., and Stammers, H. (2019). Running to Breaking Point? The Relationship Between 1.5-Mile Run Time and Injury Risk in female Recruits During British Army Basic Training. Journal of the Royal Army Medical Corps, jramc-2018.

[65] Canham, M.L., Mcferren, M.A., and Jones, B.H. (1996). The Association of Injuries with Physical Fitness Among Men and Women in Gender Integrated Basic Combat Training Units. MSMR, 2(8-10), 2.

[66] Chasse, E., Lacroche, M.A., Dufour, C., Leclerc, G., Guimond, R., and Lalonde, F. (2018). Association Between Musculoskeletal Injuries and the Canadian Armed Forces Physical Employment Standard in Canadian Military Recruits. Journal of Science in Medicine and Sport (in press).

[67] Kenny, G.P., Groeller, H., Mcginn, R., and Flouris, A.D. (2016). Age, Human Performance, and Physical Employment Standards. Applied Physiology, Nutrition, and Metabolism, 41(6), S92-S107.

[68] Jackson, A.S., Sui, X., Hébert, J.R., Church, T.S., and Blair, S.N. (2009). Role of Lifestyle and Aging on the Longitudinal Change in Cardiorespiratory Fitness. Archives of Internal Medicine, 169(19), 1781-1787.

[69] Kusy, K., and Zieliński, J. (2014). Aerobic Capacity in Speed-Power Athletes aged 20 – 90 years vs Endurance Runners and Untrained Participants. Scandinavian Journal of Medicine & Science in Sports, 24(1), 68-79.

[70] Larsson, L., Grimby, G., and Karlsson, J. (1979). Muscle Strength and Speed of Movement in Relation to Age and Muscle Morphology. Journal of Applied Physiology, 46(3), 451-456.

[71] Seiler, K.S., Spirduso, W.W., and Martin, J.C. (1998). Gender Differences in Rowing Performance and Power with Aging. Medicine and Science in Sports and Exercise, 30(1), 121-127.

[72] Martin, R.C., Grier, T., Canham-Chervak, M., Anderson, M.K., Bushman, T.T., Degroot, D.W., and Jones, B.H. (2016). Validity of Self-Reported Physical Fitness and Body Mass Index in a Military Population. The Journal of Strength & Conditioning Research, 30(1), 26-32.

[73] O'Leary, T.J., Saunders, S.C., McGuire, S.J., Venables, M.C., and Izard, R.M. (2018). Sex Differences in Training Loads during British Army Basic Training. Medicine and Science in Sports and Exercise, 50(12), 2565-2574.

[74] Moran, D.S., Evans, R., Arbel, Y., Luria, O., Hadid, A., Yanovich, R., and Finestone, A.S. (2013). Physical and Psychological Stressors Linked with Stress Fractures in Recruit Training. Scandinavian Journal of Medicine & Science in Sports, 23(4), 443-450.

[75] Richmond, V.L., Horner, F.E., Wilkinson, D.M., Rayson, M.P., Wright, A., and Izard, R. (2014). Energy Balance and Physical Demands During an 8-Week Arduous Military Training Course. Military Medicine, 179(4), 421-427.

[76] Bullock, S.H., Jones, B.H., Gilchrist, J., and Marshall, S.W. (2010). Prevention of Physical Training – Related Injuries: Recommendations for the Military and other Active Populations Based On Expedited Systematic Reviews. American Journal of Preventive Medicine, 38(1), S156-S181.

[77] Wardle, S.L., and Greeves, J.P. (2017). Mitigating the Risk of Musculoskeletal Injury: A Systematic Review of the Most Effective Injury Prevention Strategies For Military Personnel. Journal of Science and Medicine in Sport, 20, S3-S10.

[78] Knuttgen, H.G., Nordesjö, L.O., Ollander, B., and Saltin, B. (1973). Physical Conditioning Through Interval Training with Young Male Adults. Medicine and Science in Sports, 5(4), 220-226.

[79] Westcott, W.L., Skaggs, J.M., Gibson, J.R., Annesi, J.J., Reynolds, R.D., and O'Dell, J.P. (2007). Comparison of Two Exercise Protocols on Fitness Score Improvement in Poorly Conditioned Air Force Personnel. Perceptual and Motor Skills, 104(2), 629-636.

[80] Lundberg, T.R., Fernandez-Gonzalo, R., Gustafsson, T., and Tesch, P.A. (2012). Aerobic Exercise Does Not Compromise Muscle Hypertrophy Response to Short-Term Resistance Training. Journal of Applied Physiology, 114(1), 81-89.

[81] Lundberg, T.R., Fernandez-Gonzalo, R., and Tesch, P.A. (2014). Exercise-Induced AMPK Activation Does Not Interfere with Muscle Hypertrophy in Response to Resistance Training in Men. Journal of Applied Physiology, 116(6), 611-620.

[82] Williams, A.G., Rayson, M.P., and Jones, D.A. (2002). Resistance Training and the Enhancement of the Gains in Material-Handling Ability and Physical Fitness of British Army Recruits During Basic Training. Ergonomics, 45(4), 267-279.

[83] Knapik, J.J., Harman, E.A., Steelman, R.A., and Graham, B.S. (2012). A Systematic Review of the Effects of Physical Training on Load Carriage Performance. The Journal of Strength & Conditioning Research, 26(2), 585-597.

[84] Paavolainen, L., Hakkinen, K., Hamalainen, I., Nummela, A., and Rusko, H. (1999). Explosive-Strength Training Improves 5-km Running Time by improving Running Economy and Muscle Power. Journal of Applied Physiology, 86(5), 1527-1533.

[85] Rønnestad, B.R., and Mujika, I. (2014). Optimizing Strength Training for Running and Cycling Endurance Performance: A Review. Scandinavian Journal of Medicine & Science in Sports, 24(4), 603-612.

[86] Drain, J.R., Burley, S.D., Nindl, B.C., Sampson, J., and Groeller, H. (2016, May). Physical and Physiological Changes in Army Recruits in Response to Alternate Physical Conditioning Regimen. Medicine and Science in Sports and Exercise, 48, 5, 269-269. Lippincott Williams & Wilkins, Philadelphia, PA, US.

[87] Drain, J.R., Groeller, H., Burley, S.D., and Nindl, B.C. (2017). Hormonal Response Patterns are Differentially Influenced by Physical Conditioning Programs During Basic Military Training. Journal of Science and Medicine in Sport, 20, S98-S103.

[88] Carstairs, G.L., Ham, D.J., Savage, R.J., Best, S.A., Beck, B., and Billing, D.C. (2018). A Method for Developing Organisation-Wide Manual Handling Based Physical Employment Standards in a Military Context. Journal of Science and Medicine in Sport, 21(11), 1162-1167.

[89] Groeller, H., Burley, S., Orchard, P., Sampson, J.A., Billing, D.C., and Linnane, D. (2015). How Effective is Initial Military-Specific Training in the Development of Physical Performance of Soldiers? The Journal of Strength & Conditioning Research, 29, S158-S162.

[90] Rudzki, S.J. (1997). Injuries in Australian Army Recruits. Part I: Decreased Incidence and Severity of Injury Seen with Reduced Running Distance. Military Medicine, 162(7), 472-476.

[91] Knapik, J.J., Reynolds, K.L., and Harman, E. (2004). Soldier Load Carriage: Historical, Physiological, Biomechanical, and Medical Aspects. Military Medicine, 169(1), 45-56.

[92] Roy, T.C., Lopez, H.P., and Piva, S.R. (2013). Loads Worn by Soldiers Predict Episodes of Low Back Pain During Deployment to Afghanistan. Spine, 38(15), 1310-1317.

[93] Laroche, M. (2018). Comparison of Force Test Results and Injuries at Canadian Forces Leadership and Recruit School. Canadian Institute of Military Veterans Health Research, Regina, SK, Canada.

[94] Sharp, M.A., Foulis, S.A., Redmond, J.E., Canino, M.C., Cohen, B.S., Hauret, K., ... and Gebhardt, D.L. (2018). Longitudinal Validation of the Occupational Physical Assessment Test (OPAT) (No. T18-05). US Army Research Institute of Environmental Medicine Natick United States.

[95] Billing, D.C., and Drain, J.R. (2017). International Congress on Soldiers' Physical Performance 2017: Research Priorities Across the service Members Operational Lifecycle. Journal of Science and Medicine in sport, 20, S1-S3.

[96] Jones, S.B., Knapik, J.J., Sharp, M.A., Darakjy, S., and Jones, B.H. (2007). The Validity of Self-Reported Physical Fitness Test Scores. Military Medicine, 172(2), 115-120.

[97] Schuh-Renner, A., Canham-Chervak, M., Grier, T.L., and Jones, B.H. (2019). Accuracy of Self-Reported Injuries Compared to Medical Record data. Musculoskeletal Science and Practice, 39, 39-44.

CHAPTER 7 – SEX DIFFERENCES IN THE PHYSIOLOGICAL RESPONSES TO PROLONGED MILITARY WORK: IMPLICATIONS FOR PES

H.K. Teien and O. Vikmoen
Norwegian Defence Research Establishment (FFI)
NORWAY

S. Blacker
University of Chichester
UNITED KINGDOM

T. O'Leary
UK Ministry of Defence
UNITED KINGDOM

T.J. Reilly
Department of National Defence
CANADA

J. Drain
Defence Science and Technology Group
AUSTRALIA

A. Malgoyre
Institut de Recherche Biomédicale des Armées
FRANCE

H.A. Hasselstrom
Danish Defence Medical Command
DENMARK

H.C. Tingelstad
Department of National Defence
CANADA

K. Friedl
US Army Research Institute for Environmental Medicine
UNITED STATES

7.1 INTRODUCTION

Physical Employment Standards (PES) are typically performed in a controlled environment. However, the actual military tasks that PES attempts to replicate are typically performed in a multi-stressor environment, such as during sustained Military Field Exercises (MFEs). MFEs typically include inadequate sleep, high physical activity levels, inadequate energy intake, psychological strain, and high-risk environments that are rapidly changing [1]. They are also likely to involve a wide range of environmental stressors. PES primarily evaluates acute (< 4 hours) physiological performance, and does not directly test the ability to withstand all the military stressors over several days during MFEs [2]. These stressors provide challenges to individual performance capabilities that are not typically evaluated as part of PES testing. A relatively simple example of modelling the metabolic cost and speed of load carriage highlights the complexity of trying to account for these variables to accurately predict real world job performance. These models attempt to model the metabolic cost and how fast individuals, of different body size, can walk when carrying varying rucksack loads over different surfaces for varying durations. These are important calculations for a mission planner to know what is feasible in a military operation, but only provide estimations that are limited by their ability to account for individual physiological responses and environmental conditions. This example is for a relatively simple task of walking with load, whereas PES often encompasses a range of physical actions (e.g., material manual handling, moving simulated casualties, intermittent sprints). Training and selection might also require this as a standard in the military job (e.g., job specialty badges in the US Army have typically required completion of a 19.3 km march with a 15.9 kg load in less than 180 minutes to quality). This performance is greatly affected by other factors that have not been successfully added into the model such as performing the same task in hot or cold conditions, or at altitude. Even more relevant to testing for military mission capabilities might be the effects of limited energy intake and sleep. Thus, it is not a trivial challenge to adjust PES to likely realistic operational demands and only the most important variables can be reasonably considered at this time. One of those is sex, including how men and women differentially respond to operational stressors. The additional demands on performance observed during MFEs have been only partially quantified primarily for men and even less information is available for women. Nevertheless, sex differences in performance response to these stressors have been postulated and/or actually observed. The differential performance decrements produced by various MFE stressors in men and women are presented in this section.

High physical activity levels with restricted food intake produces a negative energy balance and may result in a loss in body mass [3], [4], [5], impaired aerobic capacity [6], anaerobic capacity [7], and muscle strength [8], [9]. This negative energy balance induces compensatory physiological and behavioural mechanisms such as lower resting metabolic rate and increased mechanical efficiency [10], and altered immune and endocrine function [11], [12], [13], [14]. Nearly all studies in this area have been conducted with men, however new data is now emerging as the proportion of women increases in most national militaries and they are more likely to conduct this type and intensity of military training [15], [16].

7.2 EFFECTS ON BODY MASS AND BODY COMPOSITION

Women are better able to maintain lean body mass during strenuous MFEs with food deprivation compared to men, and this is attributable to lower absolute energy demands for movement (due to typically smaller body size) and an estrogen-dependent enhanced fat oxidation which is glycogen and protein sparing [15], [17]. Hoyt et al. [15] compared changes in body mass and body composition of men and women undergoing 7 days of sustained MFE and food deprivation. Both the men and the women lost body mass, Fat Mass (FM) and Lean Body Mass (LBM). The relative reduction in LBM was similar in men and women (~6%), while the absolute reduction in LBM (2.6 ± 1.1 vs 4.0 ± 1.2 kg) and total body mass (6.0 ± 1.1 vs 7.5 ± 1.3 kg) was smaller in women than men. There were no differences in relative energy expenditure between men and women (343 ± 26 vs 354 ± 18 kJ·kg body $wt^{-1}·day^{-1}$, respectively), however, women had a higher fat oxidation per kg of LBM compared to men (7.3 ± 0.5 vs 5.2 ± 1.0 $mg·min^{-1}·kg\ LBM^{-1}$, respectively). Lower absolute energy cost for women conducting the same field training as the men has been observed over a wide range of military exercises and is consistently explained by differences in body mass [18]. Women also gain more absolute lean mass and lose greater fat mass than men when subjected to the same basic training demands even without restricted energy intake and sleep. In this situation involving new Army recruits, it is probably explained by a lower initial physical training status in women. This is supported by data on sex difference in US sports participation. It might also be because of an inherent biological difference where estrogen favours lipid energy metabolism [Steve Foulis, unpublished results, 2019].

Raustoel [17] made similar observations when investigating changes in body mass and composition in male and female Norwegian Special Forces recruits during 5 days of selection MFE, involving strenuous physical activity (> 6000 $kcal·day^{-1}$), restricted energy intake (~700 $kcal·day^{-1}$) and inadequate sleep (1 – 6 $hours·day^{-1}$). Energy expenditure estimated from accelerometer data was higher in men compared to women (7235 ± 408 $kcal·day^{-1}$ vs 6071 ± 357 $kcal·day^{-1}$, respectively), however, when adjusted for body mass (kcal·kg body $wt^{-1}·day^{-1}$) there were no differences between men and women. Men experienced a significant loss in LBM (−2.7 ± 1.0 kg), however, LBM was preserved in women, and the reduction in absolute FM was significantly higher in the women compared to men (−2.8 ± 1.3 vs −1.8 ± 1.1 kg).

A preservation of LBM, despite large losses of body mass (~10 kg, 13%), was also recently seen by Gifford et al. [19] in six women in the British Army who completed a 61 day unassisted traverse of the Antarctic. This preservation of LBM may have contributed to largely unaffected endocrine and metabolic function despite a severe energy deficit [19], although bone health was impaired [20]. These data support the notion of a higher proportion of energy derived from fat oxidation in women under caloric restriction and an ability to preserve LBM. These data are also supported by laboratory studies, showing that women oxidize proportionately more fat and less carbohydrates at submaximal intensities than men [21], [22].

7.3 PHYSICAL PERFORMANCE, FATIGABILITY AND RECOVERY

The strenuous physical activity, inadequate food intake, and sleep deprivation associated with MFEs impairs physical performance [5], [9], [23], [24], however, few studies have investigated the decline in physical performance in women in these settings. Raustoel [17] reported reductions in vertical jump height, maximal lower body power, and medicine ball chest throw performance following a five day strenuous selection MFE in Norwegian male and female Special Forces recruits. The decrease in vertical jump height during the MFE was similar for men and women (−19 ± 7% vs −18 ± 11%, respectively), however, the recovery was faster in women compared to men at 72 hours (−24 ± 7% vs −14 ± 8%) and 2 weeks (−17 ± 6% vs −9 ± 8%) following the MFE. Although women recovered jump height faster than men, men jumped higher than women at all time points. Before the MFE, men jumped 33% higher than women, but because the faster recovery in women this difference was reduced to 22% after 2 weeks of recovery. There was also a significant decrease in performance from 24 to 72 hours post MFE in the male group that was not observed in the female group (−5.9 ± 6.2% and 2.7 ± 5.2%, respectively). A similar trend was seen for maximal lower body power during the take-off phase in the vertical jump, which declined following the MFE in both sexes, but remained significantly reduced 2 weeks after the MFE in men only (−10.6 ± 5.8% vs −3.6 ± 7.2% for men and women respectively). There were no sex differences in medicine ball throw performance and anaerobic performance, which were reduced following the MFE, and recovered within 1 week and 2 weeks, respectively. The results from this study seem to highlight a resistance to muscle fatigue or a smaller absolute reduction in LBM and shorter recovery time for jump height in women compared to men following a strenuous military MFEs. In this study men and women performed mostly the same activities throughout the MFEs, although the MFEs were held separately for men and women, and therefore might have differed in total activity exposure. The results from this study also highlight that very demanding MFEs can lead to large reductions in performance that may take a long time to recover from in both men and women.

Data gathered during acute exercise events such as loaded marches and sustained contractions support the findings from MFEs. A study by O'Leary et al. [25] investigated the effect of 9.7 km loaded marching completed in 90 min on vertical jump height and isometric Maximal Voluntary Contraction (MVC) in male and female British Army recruits. The mass carried was military trade dependent, resulting in an average load which was not significantly different between sexes (16 ± 2 vs. 15 ± 1 kg for men and women respectively). Although the reduction in vertical jump height was the same for both men and women following the march (−5 ± 11% and −5 ± 6%, respectively), the loss in isometric MVC force was significantly less in women (−9 ± 13% vs −12 ± 9%). The sex difference in MVC loss could be due to differences in skeletal muscle fatigability between men and women. Women are more fatigue resistant during static and slow velocity movements [26], [27], but this difference is diminished during higher velocity concentric contractions [28], which may explain why no difference was observed in reduction in vertical jump between sexes. The increased fatigue resistance in women during low velocity contractions could be due to the lower muscle oxygen demand, less mechanical compression of local vasculature, a lower total muscle mass and force generation when performing the same relative work [29], a higher proportion of type I muscle fibres and a higher reliance of fat oxidation [30].

The United States Air Force (USAF) performed studies on USAF Special Warfare (SW) airmen and Explosive Ordnance Disposal (EOD) airmen participating in a 33 hour USAF physically arduous MFEs in different environments. One unit conducted jungle operations and another conducted mountain operations. Physical fitness testing pre and post demonstrated declines in physical performance, and these data were used to adjust physical standards "to ensure operators are physically capable throughout a mission" [Baumgartner, unpublished data, 2019]. This strategy of adjusting to higher standards in anticipation of stress-related reductions in performance is complicated, as operational stress conditions can reach a level that would limit performance of any human (e.g., full chemical protective equipment in high heat conditions, etc.). Thus, the new adjusted standards would have to specify an adjustment for specific hot humid and altitude exposures but not hot dry, higher altitude, cold, etc. conditions.

7.4 SEX DIFFERENCES IN CARDIOVASCULAR STRAIN DURING MILITARY TRAINING

The aforementioned studies suggest that women have a lower reduction in absolute LBM, lower muscle fatigue and improved recovery compared to men following strenuous MFEs. These data must be considered in the context of the sex differences in the physiological stress during military training and MFEs. Blacker et al. [31] reported that during a nine day monitoring period of British Army basic training, that despite similar levels of daily physical activity, women in a mixed-sex platoon were working at a significantly higher percent of their Heart Rate (HR) reserve than their male counterparts, likely due to lower aerobic fitness. When training in single-sex platoons, female British Army recruits spend more time in higher HR zones, have a higher training impulse and report more daily physical fatigue and muscle soreness compared to men, despite completing lower daily distance [32]. Continual work at an elevated HR could increase fatigue and injury risk of women compared to men, which can be counteracted by well-designed training programs, ability-based training, for example in single-sex training [32], and by screening applicants with PES.

Data gathered during acute load carriage tasks support the results observed during military training. O'Leary et al. [25] reported that women had a higher HR and Rate of Perceived Exertion (RPE) compared to men during a 9.7 km loaded march carrying a trade specific external load. This can, in part, be explained by women carrying an external load equal to 29% of their body mass, whereas men carried a mass equal to only 14% of their body mass. As muscular strength is lower in women compared to men, and significant associations were observed between muscle strength and both HR and RPE during the march, improving muscular strength in women could have a beneficial effect on the physiological strain experienced during loaded marching and reduce fatigue and injury risk.

The trainability of women was intensively investigated in two major studies in the 1990s. This is an important consideration in PES because capabilities which can be readily trained might change views on initial selection criteria, just as many jobs seek individuals based on other factors such as ability to learn and work in a team and then provide specialty training after hiring. Harman et al. [33] examined the benefits of a 24 week intensive physical training program on strength performance of young women, finding that initially 24% could lift 45.4 kg to the bed of a military truck but 78% could perform this test by the end of training. One finding related to initial lean mass, where women with at least 50 kg of lean mass at the start of the study were trainable while women starting below 35 kg of lean mass were not successful [33]. This might suggest that a minimum lean mass rather than a lifting strength capability might be the more appropriate PES for initial job selection for women and, likely, for men as well. Another study, by Kraemer et al. [34], involved large groups of young women with upper or lower or both upper and lower body training compared to a control group of men and women and demonstrated that generalized and specific upper body strength training programs were more beneficial than specific lower body strength training in narrowing the differences in strength performance between men and women [34]. This makes sense as differences in lower body strength are much less between men and women and androgens have predominant effects on upper body musculature. This study provided confirmatory data for the high degree of strength trainability in healthy young women.

All physical performance data with women must be considered in the context of their micronutrient status, particularly iron status. Women have a high incidence of iron deficiency and even iron deficiency anemia during the stress of initial entry training and this poor iron status consistently deteriorates through the end of this training period [35], [36], [37]. When corrected with an iron supplement, women originally assessed with iron deficiency anemia improved their two mile run time by two minutes compared to the women not supplemented [35], [36], [37]. In the very rigorous US Army Ranger training course with inadequate energy intake and high workload, iron status of the men was not affected at any time point. Ignoring this aspect can very likely produce flawed performance comparisons between the men and women, as men are not affected in a similar manner [38].

7.5 SUMMARY

Compared to men, women appear to experience greater physiological strain for the same amount of physical activity in military settings, likely due to on average lower muscle mass and aerobic fitness. However, it appears that women experience smaller reductions in absolute LBM, lower muscle fatigue and improved recovery compared to men following strenuous MFEs. Further research should include men and women that together are exposed to a standardized MFE to further examine sex differences in response to MFEs. The new research must also use an integrative approach to the evaluation of performance in MFEs as physical performance is not an isolated component but one that is also affected by fatigue resistance and other factors that may be differently preserved in women than men. If role-related PES were performed to an individual best effort before and after MFEs, sex- or physical fitness- specific differences in role-related physical performance impairment could be determined.

7.6 REFERENCES

[1] Tharion, W.J., Lieberman, H.R., Montain, S.J., Young, A.J., Baker-Fulco, C.J., Delany, J.P., and Hoyt, R.W. (2005). Energy Requirements of Military Personnel. Appetite, 44(1), 47-65.

[2] Nindl, B.C., Billing, D.C., Drain, J.R., Beckner, M.E., Greeves, J., Groeller, H., ... and Taylor, N.A. (2018). Perspectives on Resilience for Military Readiness and Preparedness: Report of an International Military Physiology Roundtable. Journal of Science and Medicine in Sport, 21(11), 1116-1124.

[3] Opstad, P.K. (1995). Medical Consequences in Young Men of Prolonged Physical Stress with Sleep and Energy Deficiency (No. NDRE/PUBLICATION-95/05586). Norwegian Defence Research Establishment Kjeller.

[4] Friedl, K.E., Moore, R.J., Martinez-Lopez, L.E., Vogel, J.A., Askew, E.W., Marchitelli, L.J., ... and Gordon, C.C. (1994). Lower Limit of Body Fat in Healthy Active Men. Journal of Applied Physiology, 77(2), 933-940.

[5] Nindl, B.C., Barnes, B.R., Alemany, J.A., Frykman, P.N., Shippee, R.L., and Friedl, K.E. (2007). Physiological Consequences of US Army Ranger Training. Medicine & Science in Sports & Exercise, 39(8), 1380-1387.

[6] Guezennec, C.Y., Satabin, P., Legrand, H., and Bigard, A.X. (1994). Physical Performance and Metabolic Changes Induced by Combined Prolonged Exercise and Different Energy Intakes in Humans. European Journal of Applied Physiology and Occupational Physiology, 68(6), 525-530.

[7] Hackney, A.C., Shaw, J.M., Hodgdon, J.A., Coyne, J.T., and Kelleher, D.L. (1991). Cold Exposure During Military Operations: Effects on Anaerobic Performance. Journal of Applied Physiology, 71(1), 125-130.

[8] Johnson, M.J., Kriedl, K.E., Frykman, P.N., and Moore, R.J. (1994). Loss of Muscle Mass is Poorly Reflected in Grip Strength Performance in Healthy Young Men. Army Research Institute of Environmental Medicine, Natick, MA, US.

[9] Knapik, J.J., Jones, B.H., Meredith, C., and Evans, W.J. (1987). Influence of a 3.5 Day Fast on Physical Performance. European Journal of Applied Physiology and Occupational Physiology, 56(4), 428-432.

[10] Bahr, R., Opstad, P.K., Medbø, J.I., and Sejersted, O.M. (1991). Strenuous Prolonged Exercise Elevates Resting Metabolic Rate and Causes Reduced Mechanical Efficiency. Acta Physiologica Scandinavica, 141(4), 555-563.

[11] Gomez-Merino, D., Chennaoui, M., Burnat, P., Drogou, C., and Guezennec, C.Y. (2003). Immune and Hormonal Changes Following Intense Military Training. Military Medicine, 168(12), 1034-1038.

[12] Bøyum, A., Wiik, P., Gustavsson, E., Veiby, O.P., Reseland, J., Haugen, A.H., and Opstad, P.K. (1996). The Effect of Strenuous Exercise, Calorie Deficiency and Sleep Deprivation on White Blood Cells, Plasma Immunoglobulins and Cytokines. Scandinavian Journal of Immunology, 43(2), 228-235.

[13] Opstad, P.K. (1991). Alterations in the Morning Plasma Levels of Hormones and the Endocrine Responses to Bicycle Exercise During Prolonged Strain. The Significance of Energy and Sleep Deprivation. European Journal of Endocrinology, 125(1), 14-22.

[14] Nindl, B.C., Castellani, J.W., Young, A.J., Patton, J.F., Khosravi, M.J., Diamandi, A., and Montain, S.J. (2003). Differential Responses of IGF-I Molecular Complexes to Military Operational Field Training. Journal of Applied Physiology, 95(3), 1083-1089.

[15] Hoyt, R.W., Opstad, P.K., Haugen, A.H., Delany, J.P., Cymerman, A., and Friedl, K.E. (2006). Negative Energy Balance in Male and Female Rangers: Effects of 7 d of Sustained Exercise and Food Deprivation. The American Journal of Clinical Nutrition, 83(5), 1068-1075.

[16] Westphal, K.A., Friedl, K.E., Sharp, M.A., King, N., and Kramer, T.R. (1995). Health, Performance, and Nutritional Status of US Army Women during Basic Combat Training (No. USARIEM-T96-2). Army Research Institute of Environmental Medicine, Natick, MA, US.

[17] Raustøl, M.L. (2018). Sex Differences in Effect and Recovery of Strenuous Military Field Exercises Master's thesis, Norges idrettshøgskole.

[18] Friedl, K.E., and Ness, J.W. (2008). Upper Limits of Energy Expenditure for Men and Women in Military Operational Settings: Does Female Fat Metabolism Provide a Performance Advantage. In NATO HFM Symposium-Impacts of Gender Differences on Conducting Operational Activities.

[19] Gifford, R.M., O'Leary, T., Cobb, R., Blackadder-Weinstein, J., Double, R., Wardle, S.L., ... and Wilson, A. (2019). Female Reproductive, Adrenal, and Metabolic Changes during an Antarctic Traverse. Medicine and Science in Sports and Exercise, 51(3), 556-567.

[20] O'Leary, T.J., Gifford, R.M., Double, R.L., Reynolds, R.M., Woods, D.R., Wardle, S.L., and Greeves, J.P. (2019). Skeletal Responses to an All-Female Unassisted Antarctic Traverse. Bone, 121, 267-276.

[21] Hunter, S.K. (2016). The Relevance of Sex Differences in Performance Fatigability. Medicine and Science in Sports and Exercise, 48(11), 2247.

[22] Tarnopolsky, M.A., and Ruby, B.C. (2001). Sex Differences in Carbohydrate Metabolism. Current Opinion in Clinical Nutrition & Metabolic Care, 4(6), 521-526.

[23] Legg, S.J., and Patton, J.F. (1987). Effects of Sustained Manual Work and Partial Sleep Deprivation on Muscular Strength and Endurance. European Journal of Applied Physiology And Occupational Physiology, 56(1), 64-68.

[24] Montain, S.J., and Young, A.J. (2003). Diet and Physical Performance☆ In Special Section on US Army Research, Guest-edited by M.L. Meiselman, F.M. Kramer, J. Soyer and P. Pliner. Appetite, 40(3), 255-267.

[25] O'Leary, T.J., Saunders, S.C., McGuire, S.J., and Izard, R.M. (2018). Sex Differences in Neuromuscular Fatigability in Response to Load Carriage in the Field in British Army Recruits. Journal of Science and Medicine in Sport, 21(6), 591-595.

[26] Bam, J., Noakes, T.D., Juritz, J., and Dennis, S.C. (1997). Could Women Outrun Men in Ultramarathon Races?. Medicine and Science in Sports and Exercise, 29(2), 244-247.

[27] Hunter, S.K., and Enoka, R.M. (2001). Sex Differences in the Fatigability of Arm Muscles Depends on Absolute Force During Isometric Contractions. Journal of Applied Physiology, 91(6), 2686-2694.

[28] Senefeld, J., Yoon, T., Bement, M.H., and Hunter, S.K. (2013). Fatigue and Recovery from Dynamic Contractions in Men and Women Differ for Arm and Leg Muscles. Muscle & Nerve, 48(3), 436-439.

[29] Epstein, Y., Yanovich, R., Moran, D.S., and Heled, Y. (2013). Physiological Employment Standards IV: Integration of Women in Combat Units Physiological and Medical Considerations. European Journal of Applied Physiology, 113(11), 2673-2690.

[30] Hicks, A.L., Kent-Braun, J., and Ditor, D.S. (2001). Sex Differences in Human Skeletal Muscle Fatigue. Exercise and Sport Sciences Reviews, 29(3), 109-112.

[31] Blacker, S.D., Wilkinson, D.M., and Rayson, M.P. (2009). Gender Differences in the Physical Demands of British Army Recruit Training. Military Medicine, 174(8), 811-816.

[32] O'Leary, T.J., Saunders, S.C., McGuire, S.J., Venables, M.C., and Izard, R.M. (2018). Sex Differences in Training Loads During British Army Basic Training. Medicine and Science in Sports and Exercise, 50(12), 2565-2574.

[33] Harman, E., Frykman, P., Palmer, C., Lammi, E., and Reynolds, K. (1997). Effects of a Specifically Designed Physical Conditioning Program on the Load Carriage and Lifting Performance of Female Soldiers (No. USARIEM-T98-1). Army Research Inst of Environmental Medicine, Natick, MA, US.

[34] Kraemer, W.J., Mazzetti, S.A., Nindl, B.C., Gotshalk, L.A., Volek, J.S., Bush, J.A., ... and Fleck, S.J. (2001). Effect of Resistance Training on Women's Strength/Power and Occupational Performances. Medicine & Science in Sports & Exercise, 33(6), 1011-1025.

[35] McClung, J.P., Karl, J.P., Cable, S.J., Williams, K.W., Nindl, B.C., Young, A.J., and Lieberman, H.R. (2009). Randomized, Double-Blind, Placebo-Controlled Trial of Iron Supplementation in Female Soldiers During Military Training: Effects on Iron Status, Physical Performance, and Mood. The American Journal of Clinical Nutrition, 90(1), 124-131.

[36] McClung, J.P., Karl, J.P., Cable, S.J., Williams, K.W., Young, A.J., and Lieberman, H.R. (2009). Longitudinal Decrements in Iron Status During Military Training in Female Soldiers. British Journal of Nutrition, 102(4), 605-609.

[37] McClung, J.P., Marchitelli, L.J., Friedl, K.E., and Young, A.J. (2006). Prevalence of Iron Deficiency and Iron Deficiency Anemia Among Three Populations of Female Military Personnel in the US Army. Journal of the American College of Nutrition, 25(1), 64-69.

[38] Moore, R.J., Friedl, K.E., Tulley, R.T., and Askew, E.W. (1993). Maintenance of Iron Status in Healthy Men During an Extended Period of Stress and Physical Activity. The American Journal of Clinical Nutrition, 58(6), 923-927.

CHAPTER 8 – RETURN-TO-DUTY PHYSICAL RECOVERY TIMELINES AND POLICIES, INCLUDING POST-PARTUM, POST INJURY, AND AMPUTEES

A. Fieldhouse
UK Ministry of Defence
UNITED KINGDOM

S.A. Foulis and M. A. Sharp
US Army Research Institute
of Environmental Medicine
UNITED STATES

T.J. Reilly
Department of National Defence
CANADA

H. Kilding
New Zealand Defense Force
NEW ZEALAND

S. Blacker
University of Chichester
UNITED KINGDOM

K.E. Friedl
US Army Research Institute
of Environmental Medicine
UNITED STATES

8.1 RETURN-TO-DUTY STANDARDS

Decrements in physical performance capacity typically occur following pregnancy and major injury or illness. Return to duty requires policies which consider specialized physical training and strategies, including timing, for safe physical testing to determine capabilities relative to PES. Earlier policies were quick to retire a soldier following a significant state change while most armies today recognise the capability for recovery with scientifically based training programs, the willingness of soldiers to return to full participation, and the need to retain experienced performers. This chapter considers return to duty PES for three categories of soldiers where there may be a temporary incapacity: post-partum women; recovered from serious injury or illness; and rehabilitated amputees.

8.1.1 Pregnancy and Post-Partum

8.1.1.1 Introduction

Women account for 9.4% of the British Army, and 95% are of childbearing age [1]. Based on the age demographic of the US Army, approximately 7 – 9% of US female soldiers will be pregnant at any given time. The American College of Obstetrics and Gynaecology (ACOG) "Committee on Employment Considerations during Pregnancy and the Post-partum Period" cites US census data for the US civilian workforce which indicates that "56% of women worked full time through pregnancy, with 82% of nulliparous women continuing to work to within one month before their due dates" [2]. Women are now allowed to continue their military service during pregnancy and following maternity leave, and societal changes, legal protection of women against discrimination, and financially supportive maternity packages mean pregnant soldiers and working military mothers are a prevalent and normal part of a modern military. There is a growing understanding of the short and long-term health effects of pregnancy on a servicewoman's health, and its impact on her ability to achieve full occupational fitness and meet Physical Employment Standards (PES). Pre-conception education of women is required to ensure that they fully understand the risks to themselves and their offspring, particularly if they should choose not to disclose their pregnancy. It is recommended that medical restrictions are intelligently applied whenever pregnancy is detected, without disclosure of the underlying reason, even if the woman has no intent to continue a pregnancy to term, because the risks to her own health remain extant.

8.1.1.2 Considerations for Military Employment During Pregnancy

8.1.1.2.1 Exercise in Pregnancy

There is a relatively large body of experimental data demonstrating the resilience of human pregnancy to cardiovascular strain associated with work. An ovine exercise model provided insights into the buffering action of the placenta in cardiovascular dynamics and metabolism between mother and developing foetus during exercise, supported by new data on exercising women [3]. This physiological resilience of pregnancy supports guidelines for reasonable work levels through normal pregnancy [4]. However, during pregnancy, there is a progressive increase in resting oxygen consumption and an increased energy demand for respiration as the growing foetus puts pressure on the diaphragm, with a net decline in exercise related aerobic capacity [4].

Beneficial maternal outcomes associated with exercise through pregnancy include lower fat accretion and better maintenance of aerobic fitness [5]. An effect of exercise on birthweight has not been well demonstrated, where slightly lower average birthweights of babies born to exercising mothers is confounded by energy intake [4]. From animal studies and only limited human data, there appears to be an increased risk of congenital malformations with exposure to hyperthermia (> 39 C) during pregnancy [6]. Human data on outcomes associated with exposure to exercise-induced hyperthermia is limited: a similar rise in core temperature might occur in elite marathon running performance, but there is a growing body of evidence on safe pregnancy outcomes for female distance runners continuing their training and even participating in races to full term pregnancy [7].

8.1.1.2.2 Biomechanical Changes in Pregnancy

During pregnancy there is an increased risk of falling and musculoskeletal injuries. Approximately 25% of pregnancies involve a fall [8], often due to poor dynamic postural balance. Decrements are mainly observed during the third trimester of pregnancy compared with the other pregnancy trimesters and non-pregnant women [9]. This may reflect the effects of weight gain and the shift in the centre of mass during pregnancy. However, an investigation with self-controlled subjects indicated that the decreased postural stability in the third trimester was potentially related to laxity of the pelvic ligaments, rather than the increase in the height of the centre of gravity and the increase in weight because of the developing foetal load [10].

Many of the biomechanical changes during pregnancy occur in the pelvis or in adjacent joints. The position taken by the pelvis in the third trimester seems to be a consequence of the weight of the uterus, placenta, and foetus being placed on the anterior zone of the body, and the weaker capacity to produce force by the rectus abdominis [11]. These changes bring consequences for the muscles attached to the pelvis, and particularly a greater strain on the abductors and extensors muscles of the hip, which, combined with a higher stretch derived from the anterior tilt of the pelvis, will contribute to lower back, pelvis, hip, and sacroiliac pain [12], [13]. These types of pain are also associated with a decreased activation of knee extensors and ankle plantar-flexors [14] which result in a modified gait pattern. These changes in gait pattern may contribute to the higher probability of trips and falls [15].

8.1.1.2.3 Nutrition and Requirement for Adequate Weight Gain During Pregnancy

Overly stringent occupational standards for weight/body fat for military women may inadvertently cause behaviours during pregnancy that harm the health of the offspring. An important observation in a cohort of Dutch Army recruits whose mothers had been semi-starved during pregnancy in the Dutch "Hunger Winter" led to the discovery of important epigenetic effects associated with the nutritional status of pregnant women [16]. Women underfed early in pregnancy produced hyperphagic offspring with substantially increased risk of obesity; women underfed in the second half of pregnancy produced underweight offspring with other later health consequences and lower adult height attainment [17]. Subsequent animal and human studies have confirmed these findings. Very recent findings that the Japanese population is losing height, with individuals born after 1980 being shorter than previous era cohorts (who had steadily gained in stature with improved nutrition), has been ascribed to current cultural influences that favour slimness in women, apparently extending to maintaining a slim appearance during late pregnancy [18]. Adequate weight gain must therefore be encouraged during pregnancy, with education and support to help prevent unhealthy lifestyle choices and minimise their impact: short-term on pregnancy outcomes, mid-term on return to occupational fitness and long-term on health outcomes for both child and mother.

Unhealthy lifestyles will contribute to deconditioning, with implications for the return to occupational fitness post-partum, but an overly stringent focus on weight/body fat could cause behaviours during pregnancy that harm the health of the offspring. In both the US and British Army, body composition standards and physical performance tests are waived from diagnosis of pregnancy, until a variable period post-partum. For the US, these women are given six months after the end of pregnancy. For the UK, women are given six months after return to work and up to 18 months after delivery. In the US guidance on physical training for pregnant and post-partum soldiers has also been developed to assist in the post-partum transition back to fitness and body composition standards [19], as well as to reduce stress in pregnancy, and the UK is introducing enhanced physical development and health education support to pregnant and post-partum women.

US Army-sponsored research conducted by Nancy Butte and colleagues quantified food energy requirements during pregnancy, helping to establish new national guidelines for appropriate weight gain to sustain optimal foetal and maternal health outcomes [20]. Previously, the energy requirements of pregnancy had been underestimated. Appropriate pregnancy weight gain depends partly on the initial nutritional status of the mother at the start of pregnancy, with adjusted recommendations for initially overweight and underweight women [21].

8.1.1.2.4 Workplace Hazards for Pregnant Women

Military workplace hazards can be categorised as chemical, physical, biological and psychosocial. Chemical hazards include exposure to toxins with implications for foetal development, such as airborne lead on firing ranges, permethrin in impregnated uniforms and petroleum products used in transport roles (Figure 8-1).

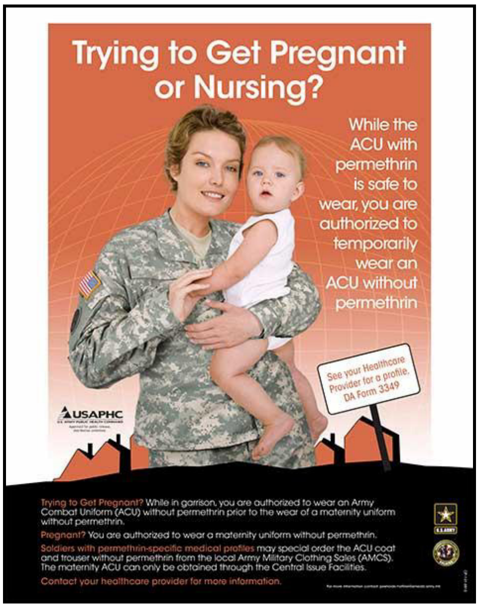


Figure 8-1: Exception Provided for Permethrin-Free Uniform Wear During Pregnancy to Prevent Unnecessary Exposure of the Developing Fetes to this Neurotoxin.

Normal workplace hazard limits (e.g., EH40/2005 Workplace Exposure Limits) are largely irrelevant for the pregnant and nursing soldiers. This is because the foetus is considered more vulnerable due to its rapid growth and development, particularly in early pregnancy and organ development. The risk profile of exposures is also different for the pregnant women for several reasons, including changes in metabolism which affect toxin absorption, physical changes which affect fit and effectiveness of personal protective equipment and physiological changes which affect immunity, cardiovascular reserve and ligament laxity [22]. Additionally the military workplace is often unpredictable by nature, and workplace controls are difficult to exercise. Physical risks include slips and falls, which could cause adverse pregnancy outcomes through a physical trauma [23] and pregnancy-associated ligament laxity [4] may increase the risk of subsequent maternal injury. Biological hazards such as influenza, Zika virus and malaria pose risks to the foetus and greater health risks to the soldier when pregnant [24]. There is also a special risk with Zika transmission from an infected military partner through unprotected sex, where there is reduced protection against Zika infection in the female reproductive tract [25]. Psychosocial risks include increased fatigue, and have implications for safety-critical work tasks. A summary of typical protective policies that apply to normal soldier pregnancies, based on current US policies [26], is given in the case study box below (Figure 8-2).

Occupational limitations while pregnant
UNITED STATES ARMY

Restrict exposure to military fuels

No weapons training in indoor firing ranges due to airborne lead concentrations and gas emissions

No work in motor pool involving painting, welding, soldering, grinding etc. where the soldier is routinely exposed to carbon monoxide, diesel exhaust, hazardous chemicals, paints, organic solvent vapours, or metal dusts and fumes

The Soldier must avoid excessive vibrations… [that] occur in larger ground vehicles… driven on unpaved surfaces

The Soldier is exempt from wearing load bearing equipment to include the web belt, individual body armour… is not recommended and must be avoided after 14 weeks gestation

The Soldier is exempt from all immunizations except influenza and tetatus-diptheria, and from exposure to all fetotoxic chemicals

The Soldier must not climb or work on ladders or scaffolding

At 20 weeks of pregnancy, the Soldier is exempt from standing at parade rest or attention for longer than 5 minutes

At 28 weeks of pregnancy, the Soldier must be provided a 15 minute rest period every 2 hours… and must not work more than 8 hours in any 1 day… including one hour for physical training

Body composition standards and physical performance tests are waived from diagnosis of pregnancy

Figure 8-2: Occupational Limitations While Pregnant, US Army.

In 2017, with more evidence to support the ability to continue working during pregnancy, the Canadian Armed Forces updated their minimum temporary medical employment limitations assigned for pregnancy in order to: Incorporate new evidence-based occupational health recommendations for maternal physical activity and foetal exposure to noise, toxins and electromagnetic radiation; and remove prescriptive restrictions "unfit field" and "unfit sea."

It is a basic tenet of CAF policies relating to the personnel management of pregnant members that the member shall not be disadvantaged in any way simply because she is pregnant (Defense Administration Orders and Directives **DAOD 5003-5**). Because a medical category may result in inability to deploy and to undertake physical fitness testing this results in the pregnant member being in technical violation of the CAF Universality of Service Principle, and therefore are not promotable. For this reason, the medical category of the member immediately prior to the diagnosis of pregnancy is to be used to determine fitness for promotion. In addition, these new revisions to DAOD 5003-5 permit some flexibility of employment in the field (maritime environment), this policy is included in detail in the case study box below (Figure 8-3). The recommendations for servicewomen in the UK follow in three separate case study boxes (Figure 8-4(a), (b), (c)).

8.1.1.3 Considerations for Military Employment Post-Partum

8.1.1.3.1 Key Factors and Emerging Science in Post-Partum RTD

The likelihood of successfully returning a soldier to full military fitness following pregnancy depends on both medical and non-medical factors. Medical factors include the duration of gestation (and whether or not the offspring survived), the mode of delivery and the presence or absence of maternal medical complications. Non-medical factors influence motivation to return to work and include legal requirements, the availability of maternity benefits and financial support, societal and workplace expectations (including the availability of childcare and attitudes towards management of common childhood illnesses) and the employer's appetite to maximise talent and exploit the benefits of diversity and working mothers within the workplace.

There is a growing understanding of the short and long-term health effects of pregnancy on a servicewoman's health, and its impact on her ability to return to full occupational fitness. Whilst some post-partum concerns are specifically attributed to complications of traumatic child-births, it is increasingly recognised that others follow apparently normal pregnancy and childbirth. These present as a continuum of signs and symptoms and include general deconditioning (due to both the direct effects of the pregnancy and the indirect effects of restricted activity during pregnancy), changes in body composition with increased ligament laxity, back and pelvic girdle pain, abdominal musculature weakness, changes in bone mass density exacerbated by lactation and pelvic floor dysfunction, with associated pelvic organ prolapse and urinary and faecal incontinence. Risk factors and the efficacy of prevention and treatment strategies are often poorly understood as the prevalence and impact of these sometimes socially embarrassing conditions is under-reported and women simply decrease performance, change or avoid all physical activity [27]. For pregnancies that do not result in a live birth, health effects are determined by the length of gestation, recognising the greater potential influence of psychological issues; the longer the gestation the greater the contribution of physical factors. The implications of these issues for returning an individual to full occupational fitness are considered below.

8.1.1.3.2 Musculoskeletal Health

Many normal pregnancies require recovery periods from joint laxity possibly related to relaxin hormone and other endocrine influences during pregnancy. The potential increased risk for musculoskeletal injury during this period is recognised to be significant but the normal time course and range of individual responses for return to pre-pregnancy muscle tension is still poorly defined. One recent review [28] suggests joints and ligaments may require up to 3 months to return to pre-pregnancy configuration.

Occupational limitations while pregnant
CANADIAN ARMED FORCES DAOD 5003-5

Following a diagnosis of pregnancy, the member will be assigned a temporary medical category with an occupational factor of O4 for a period of 12 months. The need for occupational MELs are due to all or a combination of the following reasons:

a. Ergonomics - body configuration and lack of mobility may preclude certain activities such as lying in a prone position or operating certain equipment;

b. Normal physiological changes that occur in pregnancy. These include changes that affect the CAF member's tolerance to physical activities, her balance, and her ability to safely perform certain aspects of her normal duties due to the musculoskeletal changes that are occurring, in addition to fatigue, nausea, and an increased risk of thrombosis; and

c. Hazardous exposures or activities that pose a potential risk to the member or the developing fetus, such as exposure to noise, toxins, or electromagnetic radiation

Additional specific MELs may also be required in cases of High Risk Pregnancy or in the presence of medical complications (e.g. hypertension, bleeding, or multiple pregnancy).

While the CAF recognizes the individual's autonomous right to decide what activities they can tolerate during their pregnancy, occupational MELs shall include the following, and may include other MELs as deemed appropriate by the HCP in order to lessen the risk to the individual and her pregnancy as a result of the unique stressors associated with working in a military environment:

a. Must not be employed in a position where they will routinely be exposed to loud noise (e.g. firing ranges, explosions, or work environments with noise levels > 85 dB averaged over eight working hours) (Ref E)

b. PT limited in type, duration, intensity, or frequency (Ref B);

c. Unable to do unarmed combat, collision and contact sports (Ref B);

d. Unfit military common tasks fitness evaluation (Ref B);

e. Unfit forced/ruck marching (Ref B);

f. Lifting as tolerated with the following restrictions (Ref F);

i. Until [insert date] (Less than 20 weeks of gestation)

ii. Unable to lift, push or hold items more than 16 kg on any occasion; and

iii. Unable to repetitively lift more than 14kg (<1 hour/day) or 8 kg (≥1 hour/day).

iv. After [insert date] (Greater than 20 weeks gestation)

v. Unable to lift, push or hold items more than 12 kg on any occasion: and

vi. Unable to repetitively lift more than 10 kg (<1 hour/day) or 6 kg (≥1 hour/day).

vii. Unable to perform drill and parades for greater than 30 minutes (Ref B);

viii. May require frequent rest and the opportunity to change position as needed (Ref B);

ix. Requires regularly scheduled meals;

x. Unfit firing of personal weapons and handling of lead-based ammunition (Ref E);

xi. Unfit climbing ropes or rope ladders;

xii. Unfit climbing solid ladders that are not firmly secured at the top and bottom;

xiii. Unfit exposure to substances known or suspected to be harmful to fetal development (Ref E);

xiv. Unfit exposure to ionizing radiation cumulative dose of 4 mSv for the duration of the pregnancy if member wears a dosimeter (Ref H);

xv. Unfit work with or near equipment that emits ionizing radiation if member does NOT wear a dosimeter (Ref H);

xvi. Unfit respirators unless in an emergency (Ref I);

xvii. Unfit work onboard ships at sea as of 20 weeks' gestation;

xviii. Unfit Sea States greater than 4;

xix. Unfit diving and Wet Pressurized Escape Training (WPET) (Ref E); and

xx. Unfit submarine service at sea and alongside.

Figure 8-3: Occupational Limitations While Pregnant, Canadian Armed Forces.

UK Military Recommendations to Protect Pregnant Servicewomen from Specific Occupational Hazards
UNITED KINGDOM

Physical Agents	
Ionising Radiation	Pregnant radiologists and radiographers are at a theoretical risk. However, the nationally-recommended exposure levels for pregnant women are generally one-tenth of the upper limits recommended for non-pregnant workers. This should constitute sufficient protection for the foetus. There is therefore no requirement for an MO to impose additional restrictions.
Non-Ionising Radiation - Short-Wave Equipment	MO should restrict pregnant physiotherapists from all duties involving short-wave therapeutic equipment. Pregnant servicewomen complaining of soft tissue or skeletal injuries should not be referred by MO for any treatment involving short-wave therapeutic equipment. MO should restrict pregnant servicewomen from all duties with Clansman HF or VHF sets, or any other high-frequency radio sets. The restriction should apply to both transmitters and receivers.
Non-Ionising Radiation -Visual Display Units (VDUs)	Where advice is sought from a pregnant VDU user, MO should offer reassurance that there is no substantiated risk. If the individual remains unconvinced or anxious, the MO should agree to restrict work with VDUs.
Noise – Vehicles and Weapons	As a sensible precaution, MO should restrict pregnant servicewomen from any travel in tracked vehicles. The same exclusion should apply to any travel (unless of only a few minutes' duration) in rotary wing aircraft, ie helicopters. As a sensible precaution, MO should restrict pregnant servicewomen from all exposure to gunfire noise. Therefore: Pregnant servicewomen should not be armed. They should not take part in any range duties, nor any military exercise where they are likely to be exposed at close range to small arms noise, heavy weapons noise, or pyrotechnics noise.
Vibration – Whole-Body and Hand-Transmitted	As sensible precaution, MO should impose the following restrictions on the employability of pregnant servicewomen: No off-road travel in military vehicles. No usage of fork lift trucks only limited travel (no more than a few minutes duration) in rotary wing aircraft, ie helicopters. Based on a detailed work history, MO should restrict prolonged usage in pregnancy of: Pneumatic or electric power tools (eg drilling machines, power saws, grinders, chipping hammers). Vibrating work pieces (eg mobile generators, compressors, pumps).
Heavy Lifting	MO should restrict all duties involving heavy lifting (eg movement of stores, erection of tentage, casualty handling). This is likely to be a hazard in many trades.
Long/ Irregular Hours of Work	MO should consider restricting work where there is a likelihood of a pregnant servicewoman having to undertake particularly long and irregular hours of work.
Night Work	As a sensible precaution, MO should restrict all night duties where the pregnant servicewoman complains of excessive fatigue resulting from night work.

Figure 8-4(a): UK Military Recommendations to Protect Pregnant Servicewomen from Specific Occupational Hazards.

UK Military Recommendations to Protect Pregnant Servicewomen from Specific Occupational Hazards

UNITED KINGDOM continued

Physical Agents Continued	
Physical Exercise	MO should not restrict normal PT or adventurous training in a pregnant servicewoman, unless there are clear contraindications to physical exercise. These contraindications include: acute infectious disease, multiple pregnancy, incompetent cervix, intrauterine growth retardation, hypertension, uterine bleeding, ruptured membranes. Pregnant servicewomen should not undertake fitness tests.
Trauma	MO should restrict all sports in all pregnant servicewomen after the first trimester. Military parachuting must not be undertaken at any stage of pregnancy. MO should advise pregnant servicewomen who work in equine divisions (eg RAVC and RMP personnel) to avoid all contact with horses on account of possible trauma. If this is impossible, the MO should consider imposing a formal restriction.
Extremes Of Heat	MO should advise pregnant servicewomen to exercise during the cool part of the day, and to ensure adequate hydration at all times. Pregnant servicewomen must not undertake CBRN training, other than in CBRN Dress Category Zero or CBRN Dress Category 1.
Extremes Of Cold	MO should advise pregnant servicewomen of the theoretical risk. They should not undertake any adventurous training which might entail prolonged exposure to extreme cold. During exceptionally cold weather (eg in Germany, Norway) pregnant servicewomen should be excused guard duty.
Electrical Contact	MO must assess the risk realistically. In most military employments, and with most electrical equipments, there is likely to be no danger at all to the pregnant servicewoman. If a known danger of electrical hazard from old or unreliable military equipment (as eg from some armoured vehicle power packs) exists, the MO should restrict pregnant servicewomen from all contact with such equipment.

Figure 8-4(b): UK Military Recommendations to Protect Pregnant Servicewomen from Specific Occupational Hazards Continued.

Pelvic girdle and low back pain are commonly associated with pregnancy and the post-partum period, and are often attributed to the effects of joint laxity, although this aetiology is not confirmed [29]. Pre-pregnancy fitness is not protective, and strenuous work, along with previous low back pain are major risk factors [29], [30]. Low back pain is common amongst military personnel and arduous military work may therefore place military mothers at increased risk of these complaints [31].

Diastis Recti Abdomini (DRA) of some degree is common in both pregnancy and post-partum and there are no clear risk factors, although a higher prevalence has been observed in women performing heavy lifting more than 20 times per week [32]. A higher prevalence may therefore be expected in military females regularly performing arduous tasks. Abdominal muscle function has been observed to improve by 6 months post-partum but not recovered to pre-partum levels [33].

Bone loss although not permanent, begins during pregnancy [34]. Bone loss is variable in different sites with different implications (femoral neck, lumbar spine) but the range of loss is typically reported as ~5 – 8%, and continues during lactation, regardless of calcium intake [35], [36]. Studies reviewing the effect of exercise suggest it has no impact on early post-partum lactation induced bone loss [37], although a small study comparing the effect of weight-bearing aerobic and resistance exercise between 4 and 20 weeks post-partum indicates no significant difference in total body and hip bone mass density but that less change occurs in lumbar spine bone mass density with exercise [38]. In the long term lactation is considered to cause no harm, with bone mass returning to normal between six and twelve months after cessation of lactation [35] and studies suggest it may actually be protective for future bone health [34], [35]. This provides additional justification for encouraging military personnel to breastfeed, acknowledging the employment restrictions that may be required during, and for up to 12 months after breastfeeding.

UK Military Recommendations to Protect Pregnant Servicewomen from Specific Occupational Hazards

UNITED KINGDOM continued

Biological Agents	
Cyto-Megalo-Virus (CMV)	MO should advise hospital personnel who are pregnant to avoid contact with known CMV shedders
Toxoplasma gondii	MOs should be aware of the risk to: Pregnant veterinary healthcare workers who work in veterinary hospitals which operate on cats. Pregnant veterinary healthcare workers or any personnel who work in equine divisions (where barn cats are an essential part of the establishment). They should advise such personnel accordingly, and if necessary impose a formal restriction on any contact with cats.
Chemical Agents	
Lead	MO should restrict pregnant servicewomen from all duties within indoor firing ranges.
Benzene	MO should restrict pregnant servicewomen from any direct contact with benzene or with benzene vapour, even when wearing protective equipment. Pregnant women should not be permitted to refuel military vehicles at any time. This applies also to military drivers, who must not refuel their own vehicle if pregnant.
Carbon Monoxide	MO should restrict pregnant servicewomen from all duties in vehicle parks, other than brief visits.
Anaesthetic Gases	MO should restrict servicewomen who are pregnant from any exposure to anaesthetic gases. This applies to surgeons, anaesthetists, operating theatre nurses, operating theatre technicians, etc.
Antimitotic (Cytotoxic) Drugs	MO should restrict pregnant servicewomen healthcare workers (including doctors, nurses, pharmacists and pharmacy technicians) from handling antimitotic drugs in any form.
Antimalarial Chemoprophylaxis - Mefloquine	MO should not prescribe mefloquine to any servicewoman travelling to a malarious area, unless there is no risk at all of pregnancy (eg following a hysterectomy or sterilisation).
Pesticides	Although the majority of service-approved pesticides are likely to pose no threat at all in pregnancy, MO should nevertheless restrict pregnant servicewomen from all duties involving the use of pesticides.
CS Gas	As a sensible precaution, MO should restrict pregnant servicewomen from any exposure to CS gas (eg during CBRN training).

Figure 8-4(c): UK Military Recommendations to Protect Pregnant Servicewomen from Specific Occupational Hazards Continued.

8.1.1.3.3 Pelvic Floor Dysfunction

Urinary incontinence is a relatively common post-partum problem, but is often under-reported in athletes [27], with women changing or avoiding activity rather than seeking treatment. Similar behaviours may be expected in military populations, affecting the likelihood of a mother returning to military duty. Physical health problems, including urinary and faecal incontinence, commonly reoccur or persist in first 18 months post-partum and can persist for up to 12 years [39], [40]. If healthcare providers recognise the risk factors of forceps delivery and obesity [40], and that caesarean delivery is not protective [39], they may help identify these often under-reported problems and offer the opportunity to intervene with training programs developed to prevent incontinence [41] or assist recovery [42]. There is debate as to whether these exercises are effective beyond 6 months postpartum [41] and whether they can recover all women to pre-delivery strength [43] or prevent subsequent pelvic organ prolapse [28]. This emphasises the requirement for early intervention.

8.1.1.3.4 Post-partum Psychological Health

The post-partum soldier is subject to psychological stressors which may affect workplace performance. These issues should be recognised by treating physicians as well as policymakers and employers. Post-partum depression affects many women and in some cases is serious and debilitating [44]. Chronic sleep deficiency associated with maternal responsibilities has implications for workplace safety [45]. Psychological issues may be exacerbated by anxiety about separation from a child, due to employment and possible deployment [46]. Additionally, there may be conflicts between maternal instincts and the impact of motherhood on career.

Exercise post-partum is associated with increased energy [47] and exercise is increasingly recognised to contribute to good mental health [48]. However poor mental health may affect motivation to participate in physical activities.

8.1.1.3.5 Appropriate Weight Loss and Deconditioning

Butte characterized the normal time course for return to pre-pregnancy weight and body composition [20]. This study was performed using women from the Texas Army National Guard, who were motivated by existing Army job requirements to return to baseline body composition and fitness. Butte found that initially normal weight women (average body fat of 28%; BMI 20-26 kg/m^2) gaining 30 pounds of weight through pregnancy (which included 10 pounds of fat weight), were just (on average) achieving pre-pregnancy weight and body composition by six months postpartum [20]. Women exceeding recommended healthy weight gain retained a greater amount of weight at 6 months post-partum. Reviews have identified potential interventions to address this, with modern technologies showing promise in their ability to limit post-partum weight retention [49]. Diet and exercise interventions, using objective goals and motivational devices such as HR monitors or pedometers are considered most effective [50].

Although less well defined as a secondary goal of these studies, strength and maximal aerobic performance had declined from pre-pregnancy testing at 6 weeks post-partum and still not completely recovered at six months [51]. A study of 50 active duty soldiers followed through the post-partum period, demonstrated that only 19% achieved their pre-pregnancy physical fitness scores (on the Army APFT) at six months and three out of 52 women failed. By 11 months post-partum, 31% achieved pre-pregnancy scores and only one soldier failed the APFT. The majority of women in this study felt that one year post-partum was required to reasonably achieve pre-partum standards of physical fitness [52] which is coherent with other work that recognises post-partum changes in fitness occur over 12 months [47].

These studies helped inform the current policy of the US Department of Defense which allows a minimum time of six months for all services before testing women for physical fitness and body composition. This is generally considered the minimum defensible time for military return to duty physical standards based on documented physiological changes. Six months post-partum is also a typical minimum time encouraged for breast feeding and social bonding, with workplace accommodations encouraged to support this [53].

In contrast to the multitude of studies reviewing exercise interventions to treat pelvic floor dysfunction, evidence about the effectiveness of broad all-inclusive post-partum exercise interventions are limited, especially in populations undertaking arduous activities and employment. A review of six guidelines from five countries was conducted to describe the benefits and limitations of physical activity post-partum, and its impact on breastfeeding [54]. This review recognised a lack of specificity about which physical activity is best, but noted that healthcare providers play a critical role in encouraging women to exercise and there is no adverse effect of exercising on breastfeeding. Little information is available for athletes wanting to return-to-training and a consensus group recommended that return to sport should be considered a continuum with three distinct elements: return to participation in physical activity; return to specific sport; and, return to performance [28].

The US Army has developed a Pregnancy and Post-partum Physical Training (P3T) program, which continues to evolve and improve [19], [55]. Canada introduced a guide to fitness during and after pregnancy in the CF in 2003. In addition, fitness facilitates for CAF members allow strollers and babies in the track and field house, and the fitness staff on base/wing offer pre and post-partum fitness classes for military members. CAF members are encouraged to participate in these programs and services by the base/wing Physiotherapists and medical officers.

The UK is currently introducing a pregnancy exercise and rehabilitation pathway, and is conducting research about how best to advise and women during their maternity leave, to optimise their return to occupational fitness.

8.1.1.4 Employment Policy Considerations

Employment policies must balance the demands of the host nation's relevant legislation and organisational requirements. Militaries need to consider the requirement to sustain operational effectiveness and minimise injuries, whilst maximising talent and exploiting the benefits of women in the workforce.

In accordance with both Equality and Health and Safety Legislation [56], the British Army demands its Commanders at all levels exercise a positive and supportive attitude towards pregnancy and its aftermath [57]. This is aligned with US Army pregnancy and post-partum duty policies, defined in "Standards of Medical Fitness" (Army Regulation 40-501), which specify the intent "…to protect the fetus while ensuring productive use of the Soldier" [29]. In addition, the CAF is committed to gender integration and employment equity policies as well as a supportive work environment that enables CAF members to balance military duties and family responsibilities (DAOD 5003-5 Pregnancy Administration). Attitudes towards pregnancy and post-partum policies for women in the military workplace affect readiness and resilience of the entire unit, and are moderated by leadership [58].

The policies of the British, Canadian, and US Army reflect current expert medical guidance that "working during pregnancy is generally safe" [2] and permit female soldiers to remain in the workplace, to be provided worthwhile and meaningful employment [29], [57]. In accordance with health and safety legislation, medical restrictions are recommended to protect the unborn child from specific workplace hazards. Current military guidelines for work through pregnancy and post-pregnancy are based on recommendations from expert committees of medical specialists such as the ACOG [2], [4], [59]. Recent UK Armed Forces recommendations are detailed in Figure 8-4.

Workplace pregnancy-related policies and medical advice tend to be conservative, erring on the side of safety and increased protection against medical liability, and do not stray from the standard of care established in medical practice guidelines. Physical requirements of pregnant and post-partum women in the workplace is typically treated with great caution, because of the uncertainty about safety for both mother and foetus during pregnancy, and limited evidence about return to prenatal physical fitness following birth.

Overseas reassignments and deployments are typically prohibited for pregnant personnel, reflecting the level of employment restrictions, unpredictability of workplace exposures on operations, unsuitability for health risk mitigations such as immunisations, and the need to ensure access to adequate healthcare facilities. The implications of pregnancy for a soldier's employment therefore depend both upon their ordinary role and their requirement to deploy. For example, the employment of a soldier in a base-location with clerical responsibilities will be less affected than the employment of a deployable pilot or driver, with potential exposure to petrochemicals, vibration and noise.

Medical advice should be provided on a case by case basis for early pregnancy loss, recognising the likely psychological impact and potential anxiety about the contribution of occupational factors and implications for planned future conception. As length of gestation increases the physical and psychological implications for the mother are expected to increase. Whilst 24 weeks gestation is recognised legally in the UK as the trigger point for maternity benefits: a mother who experiences intra-uterine loss or perinatal loss at or after 24 weeks gestation is afforded full maternity rights. However, 24 weeks is a fairly arbitrary cut-off that does not acknowledge the improved survival of premature infants, or reflect maternal physiological changes. Individually tailored advice is recommended for late pregnancy losses, recognising that whilst there will not be health implications of breastfeeding, psychological issues may have greater impact.

Recognising the most common post-partum issues, US Army policy [29] provides for six weeks of convalescent leave following normal pregnancy and delivery. US soldiers are encouraged to engage in the at-home component of the Army Pregnancy and Post-partum Training (P3T) program, but are exempt from the Army Physical Fitness Test (APFT) and from weight measurements for the Army Body Composition Program (ABCP) for six months after pregnancy.

UK statutory legislation excludes all servicewoman from the workplace for 2 weeks following delivery, but most will not consider return to military service before completing their 6 month fully-paid period of ordinary maternity leave [57]. A further 6 months Additional Maternity Leave (AML) is also available with decreasing pay and benefits. A UK servicewoman is also entitled to her full annual entitlement of annual leave. This may result in a potential total of 58 weeks away from the workplace during which the individual has no requirement to visit the workplace or her employer, nor complete any PES or fitness tests. On return to work the woman is routinely downgraded for at least 6 months and is exempt physical fitness tests during this period. During this period she will be encouraged to participate in the Army PT programme, which adopts a tiered approach to improve the fitness of all personnel, regardless of the reason for any loss of fitness. All Army personnel returning to full physical activity should be assessed and provided PT advice from their PT advisor (RAPTCI). Prior to this there must be clearance to exclude specialist pregnancy-related physical issues such as abdominal muscle weakness and pelvic floor dysfunction. Research has now been commissioned to look at the most effective way to physically rehabilitate British Army women back to full fitness following childbirth.

Canadian Armed Forces members are entitled to one year maternity leave with 93% of full pay. This can begin up to 8 weeks before birth and mothers can share their leave with the father/partner who can commence Paternity leave at 18 weeks post-birth and onwards with 93% full pay up to 52 weeks post-birth. While on leave members can attend medical, dental or other administrative appointments and these are not considered a return to duty. At present, they cannot complete their PES officially until they return to work. Considerations for return to work and the physical performance continuum are summarised in Figure 8-5 There is potential for many of these unknowns to be investigated with human performance research, as well as optimal training strategies to increase operational readiness post-partum.

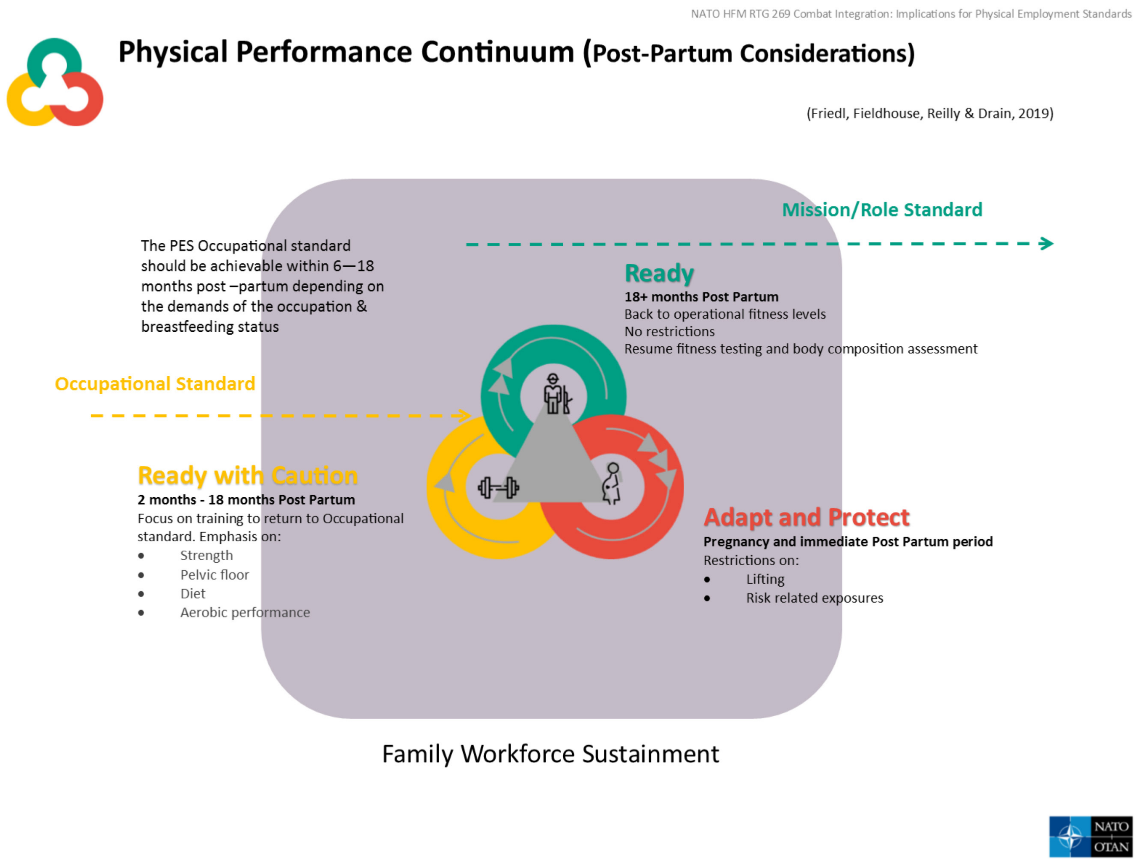


Figure 8-5: Post-Partum considerations for the Physical Performance Continuum.

8.1.1.5 Conclusions

When considering the employment of pregnant and post-partum personnel, reference must be made to the host nation's legislation, both with respect to discrimination and the health and safety requirement to protect the foetus and mother. This must be balanced with the organisational need to maximise operational effectiveness, and must therefore be personnel-led policy informed by medical advice.

8.1.2 Return to Duty Following Recovery from Illness or Injury

Returning to duty following an illness or injury is not a gender-based issue, per se. It must be acknowledged that although women make up a smaller proportion than men in most militaries, they are more likely to be injured [60]. The most commonly reported injuries in women include overuse injuries such as muscle strains, stress fractures, and sprains [61]. This trend is likely to continue as militaries open up their ground close combat roles to women. There is some evidence that when matched for body composition and physical fitness, these injury differences by gender disappear [62]. Thus, the use of properly developed PES with gender-free standards may help mitigate some of the gender differences in injury rates.

Since women are more likely to get injured, they are also disproportionately affected by programs for returning to work after injury. Many militaries provide fitness waivers for injured service members but may not provide the necessary support to aid the service member's recovery. As a result, service members may end up on these waivers (and therefore be non-deployable) for the remainder of their career. It is essential that there be a graduated pathway for injured service members and leadership to follow. This not only ensures the service member does not return to work too soon and become reinjured, but also provides clear timelines and decision points on the progress of the service member's recovery. The exact timeline will vary depending on the type and severity of the injury.

Return-to-duty decision making should be considered as part of a continuum that begins with the onset of injury, rather than an isolated step occurring at the end of medical care and rehabilitation [63]. Following an injury, there must be a progression to return the service member to full function. This can be summarised in three major phases. First, an injured service member receives acute and rehabilitative medical care to allow the injury to heal. During this time, the service member is on limited duty and is not fully participating in their unit's activities (i.e., physical training or performance of physically demanding job tasks). At the end of the rehabilitation program a medical provider (physician, physical therapist) must determine that the service member is ready to begin physical training. However, during this second phase of the continuum, there remains a period of time in which the service member may not be fully ready to train at the same intensity as the rest of the unit. The length of time to return to the pre-injury level of physical readiness is variable and dependent upon the type and severity of the injury. The goal during this phase is to ensure the injured service member has the physical capacity to both safely conduct physical training and participate in their military occupational tasks. Finally, a service member must rejoin unit physical training and pass the required PES assessment to be fully re-integrated into the unit. During each step there needs to be periodic assessments to show progress, and between each step there should be clear standards. This process is similar to the PES performance continuum discussed in Chapter 3, with 3 phases, Reset/ Readying/Ready.

The first phase in the return to work continuum is to ensure the service member is healthy enough to begin physical training. The key personnel involved in the decision making in this phase will include physicians and physical therapists. It is essential that these practitioners are familiar with the treatment and recovery of the types of injuries women are prone to (e.g., stress fractures) due to their differences in morphology and physiology. During this time period, exercise prescriptions for uninjured body regions should be provided to minimise any loss in fitness. At the end of this phase, the service member should be medically cleared to resume physical activity.

Even though a service member is cleared to resume physical training, the next phase is to ensure the service member gradually regains their fitness. Depending on the injury type, severity and rehabilitation time, it may be unreasonable to expect a service member to be ready to return to regular physical training activities with their unit. Physical training volume should be gradually increased to minimise the risk of re-injury [64]. Consultations with physical therapists and athletic trainers should be used to develop a customized physical training program to aid in preparing the service member to resume full unit activity. Checkpoints can be provided along the way to ensure the service member is making appropriate progress. During this ramp-up period, there may not be a medical professional assessing the service member's readiness to begin regular unit training. An assessment tool is needed that can be used by non-medical personnel to track the progression of recovery and estimate a service member's readiness to return to full unit training. This readiness assessment might:

- Assess a service member's physical capabilities and physical fitness/readiness to resume physical training following medical treatment and end of limited duty status;
- Include scalable and quantifiable components of muscular strength and endurance, power, agility, and cardiovascular fitness;
- Permit service members to demonstrate a progression in injury recovery and their occupational fitness/readiness relative to job-related physical demands and potential for training for the annual PES; and
- Incorporate field expedient physical performance tests and/or simulations that assess a service member's physical capabilities and are demonstrated to be predictive of performance on physically demanding occupational tasks.

The final step is for the service member to return to full operational status and pass the annual PES. A graduated increase in physical training intensity as well as physically demanding job skills practice should be provided and an appropriate timeline for full recovery must be determined to prevent re-injury. The evidence to support decisions for return to military training are scarce so we might look to the world of competitive sports for guidance [65]. Future research is needed to define, measure, and report return-to-duty and return-to-training outcomes to create valuable prognostic factors.

8.1.3 PES and its Application to Amputees and Prosthetic Evaluation

Soldiers in some countries can be retained on active duty following their recovery from serious injury and this includes permanent impairments such as limb amputations or chronic head injury. The criteria for return to duty have typically been a medical certification that the injury or illness has been treated and the individuals are back to health and can perform the physical demands of their role in their specialty.

US soldiers with a major limb amputation may request to remain on active duty. They are evaluated by a Medical Evaluation Board (MEB) for medical retention standards and then by a Physical Evaluation Board (PEB) for fitness for continued military service. They must demonstrate a K4 Medicare Functional Classification Level (i.e., prosthetic ambulation that exceeds basic ambulation skills and match those of active adult or athlete) [66]. Army amputees are generally given months of specialized rehabilitation to regain high levels of function and perform military skills.

More soldiers are remaining on active duty as amputee centres provide better prosthetics and better rehabilitation specific to their military function needs [66]. The number of amputees requesting and approved to remain on active duty increased from 11 of 469 in the 1980s (2%) to 65 of 395 in 2001 – 2006 (16%) [66]. Senior soldiers are more likely to remain on active duty, probably due to their career investment and because of their leadership roles with diminished physical demands. Various research efforts have attempted to characterize the fully capable performance of soldiers with lower limb amputations. Part of the restored capability is derived from the performance of the specific type of prosthetic leg, the extent of the amputation, and the effectiveness of the rehabilitation program [67]. In the United States Army, Soldiers with major limb loss are exempt from the body composition standards. Major limb loss is defined as an amputation above the ankle or above the wrist, which includes full hand and/or full foot loss. It does not include partial hand, foot, fingers, or toes.

In Canada, bound by the Universality of Service principle, ill and injured soldiers are required to pass the Common Military Task Fitness Evaluation (CMTFE) to remain employed as member of the Canadian Armed Forces (CAF). The PES (FORCE) is a predictive test to determine capability of a CAF member to perform the CMTFE, however as the FORCE evaluation was developed with 664 able bodied CAF members [68], and there is no evidence to suggest that this relationship, between the six task simulations in the CMTFE and its proxy predictor test, FORCE, can be applied to amputees. The CMTFE is comprised of six separate tasks: Escape to Cover, Vehicle Extrication, Stretcher Carry, Picking and Digging, Picket and Wire Carry, and Sandbag Fortification. To successfully complete the CMTFE, candidates must meet the following minimum standards: Escape to Cover (< 68 sec), Vehicle Extrication (safe completion with 86 kg), Stretcher Carry (safe completion with 86 kg), Picking and Digging (< 1080 sec per component; combined time of < 2160 sec), Picket and Wire Carry (< 1050 sec), and Sandbag Fortification (< 900 sec) [69].

The CMTFE and FORCE were implemented April 1, 2014. Subsequently, a trial was conducted with five lower extremity amputee CAF members to observe their strategies to achieve the minimum standard on the CTMFE. Participants included three above knee amputees and two below knee with prosthetics [70]. Performance outcomes were performance on the Six Common Military Tasks (time to completion), amongst other biomechanical evaluation techniques both quantitative and qualitative. For example, with regard to the picking and digging task, an able bodied CAF member will perform the task from standing and all 664 participants demonstrated a relatively similar technique when compared to that adopted by the five amputees, who employed a kneeling technique.

Following this, research to evaluate two different Lower Limb Prosthetic Devices (BiOM and X3) was conducted using the CMTFE (PES) [71]. This included an intervention crossover user-evaluation trial with repeated measures where the independent variable was the type of prosthetic device, and the main outcome measure was performance scores on the CMTFE. Secondary outcome measures included a Comprehensive High-Level Activity Mobility Predictor (CHAMP) Test, Timed Stair Ascent/Descent, and various preference surveys (Trinity Amputation and Prosthesis Experience Scales – Revised (TAPES-R), Satisfaction with Prosthesis (SAT-PRO), and the Activities-specific Balance Confidence (ABC) Scale). Results indicated for one type of prosthetic (the X3) that five of the CMTFE tasks improved (were performed faster) with the exception of "escape to cover", however this is the task for which quickness is the more operationally relevant. With the other device (the BiOM), 2 tasks improved, two showed a decline in performance, and 2 no effect.

Although in exploratory studies such as these caution should be used in drawing conclusions from these results due to the small sample size and the presence of confounding variables, the results provided valuable information to the CF Health Services team. It was recommended that a follow-up study of longer-duration be conducted to evaluate the technical performance of the trial prostheses relevant to operational demands. Utilizing the PES as an evaluation tool for prosthetic devices, demonstrates the value of an evidenced-based, task simulation PES, in lieu of a fitness predictor model for this type of research (Figure 8-6).

Figure 8-6: Researchers from the CAF Administering the PES (CMTFE) to Lower Limb Amputee.

8.1.4 Best Practices for Return to Duty After Injury

While women make up a smaller proportion than men in most militaries, they are more likely to be injured [60]. The most commonly reported injuries in women include overuse injuries such as muscle strains, stress fractures, and sprains [61]. This trend is likely to continue as militaries open up their ground close combat roles to women. There is some evidence that when matched for body composition and physical fitness, these injury differences by gender disappear [62]. Thus, the use of properly developed PESA with gender-neutral standards may help mitigate some of the gender differences in injury rates.

Since women are more likely to get injured, they are also disproportionately affected by programs for returning to work after injury. Many militaries provide fitness waivers for injured service members, but may not provide the necessary support to aid the service member's recovery. As a result, service members may end up on theses waivers (and therefore be non-deployable) for the remainder of their career. It is essential that there be a graduated pathway for injured service members and leadership to follow. This not only ensures the service member does not return to work too soon and become reinjured, but also provides clear timelines and decision points on the progress of the service member's recovery. The exact timeline will vary depending on the type and severity of the injury.

Return-to-duty decision making should be considered as part of a continuum that begins with the onset of injury, rather than an isolated step occurring at the end of medical care and rehabilitation [63]. Following an injury, there must be a progression to return the service member to full function. This can be summarised in three major phases. First, an injured service member receives acute and rehabilitative medical care to allow the injury to heal. During this time, the service member is on limited duty and is not fully participating in their unit's activities (i.e., physical training or performance of physically demanding job tasks). At the end of the rehabilitation program a medical provider (physician, physical therapist) must determine that the service member is ready to begin physical training. However, during this second phase of the continuum, there remains a period of time in which the service member may not be fully ready to train at the same intensity as the rest of the unit. The length of time to return to the pre-injury level of physical readiness is variable and dependent upon the type and severity of the injury. The goal during this phase is to ensure the injured service member has the physical capacity to both safely conduct physical training and participate in their military occupational tasks. Finally, a service member must rejoin unit physical training and pass the required PES assessment to be fully re-integrated into the unit. During each step there needs to be periodic assessments to show progress, and between each step there should be clear standards.

The first phase in the return to work continuum is to ensure the service member is healthy enough to begin physical training. The key personnel involved in the decision making in this phase will include physicians and physical therapists. It is essential that these practitioners are familiar with the treatment and recovery of the types of injuries women are prone to (e.g., stress fractures) due to their differences in morphology and physiology. During this time period, exercise prescriptions for uninjured body regions should be provided to minimise any loss in fitness. At the end of this phase, the service member should be medically cleared to resume physical activity.

Even though a service member is cleared to resume physical training, the next phase is to ensure the service member gradually regains their fitness. Depending on the injury type, severity and rehabilitation time, it may be unreasonable to expect a service member to be ready to return to regular physical training activities with their unit. Physical training volume should be gradually increased to minimise the risk of reinjury [64]. Consultations with physical therapists and athletic trainers should be used to develop a customized physical training program to aid in preparing the service member to resume full unit activity. Checkpoints can be provided along the way to ensure the service member is making appropriate progress. During this ramp-up period, there may not be a medical professional assessing the service member's readiness to begin regular unit training. An assessment tool is needed that can be used by non-medical personnel to track the progression of recovery and estimate a service member's readiness to return to full unit training. This readiness assessment might:

- Assess a service member's physical capabilities and physical fitness/readiness to resume physical training following medical treatment and end of limited duty status;
- Include scalable and quantifiable components of muscular strength and endurance, power, agility, and cardiovascular fitness;
- Permit service members to demonstrate a progression in injury recovery and their occupational fitness/readiness relative to job-related physical demands and potential for training for the annual PESA; and
- Incorporate field expedient physical performance tests and/or simulations that assess a service member's physical capabilities and are demonstrated to be predictive of performance on physically demanding occupational tasks.

The final step is for the service member to return to full operational status and pass the annual PESA. A graduated increase in physical training intensity as well as physically demanding job skills practice should be provided and an appropriate timeline for full recovery must be determined to prevent re-injury. The evidence to support decisions for return to military training are scarce so we might look to the world of competitive sports for guidance [65]. Future research is needed to define, measure, and report return-to-duty and return-to-training outcomes to create valuable prognostic factors.

8.2 THE INFLUENCE OF AEROBIC CAPACITY ON PHYSICAL TASK PERFORMANCE, INJURY RISK AND COGNITION AND WIDER OPERATIONAL EFFECTIVENESS

Chapter 5 has shown that, on average, women possess lower aerobic capacity than men. However, at an individual level, there is considerable overlap between the highest performing women and the lowest performing men.

Aerobic capacity is a critical component of the fitness required to achieve the pass standard of the various PES described in Chapter 2. The elements of PES which measure this component of fitness typically last ~5 – 120 min. However, aerobic endurance fitness is also likely to impact on wider aspects of military task performance – both physical and cognitive.

Possessing high levels of aerobic endurance fitness is important for sustained performance on prolonged military tasks and has been shown to reduce the incidence of injury during initial military training [74], [67], [76], [77] and during and following deployments [78], [79], [80]. Recent studies have also linked higher levels of aerobic endurance fitness with a reduction in fatigue during repeated sprint type activity in invasive games such as rugby and football [81], and better recovery and subsequent performance over repeated matches during tournament play [82]. Similar relationships between fatigue and recovery in military settings (e.g., whole field exercise or specific tasks such as fire and movement or digging) have not yet been established, but are possible.

Soldiers typically perform many long-term physical tasks like road marches, preparing fighting positions, filling and emplacing sandbags, constructing emplacements, loading and unloading trucks, casualty evacuation, and moving over, through and around obstacles. Individuals with higher levels of aerobic fitness perform these activities at a lower fraction, or lower percentage, of their maximal capacity (lower % $\dot{V}O_{2max}$). More aerobically fit individuals can perform tasks for longer periods of time, fatigue less rapidly and recover faster [83]. Studies have shown that as individuals fatigue, normal movement patterns change, placing more stress on parts of the body not accustomed to the stress, which could increase the likelihood of injury [84], [85], [86]. Therefore, more aerobically fit individuals are likely to have a lower risk of musculoskeletal injuries and will have greater reserve capacity for subsequent tasks. In the military setting, an individual's absolute aerobic capacity (i.e., measured in $L \cdot min^{-1}$) is likely to be as important as their aerobic capacity relative to body mass (i.e., measured in $mL \cdot kg^{-1} \cdot min^{-1}$) in influencing their physical performance.

In terms of cognition, evidence supports a positive impact of aerobic capacity on cognitive function. Hwang et al. [87] demonstrated that young adults aged 18 to 29 years (majority of the military population) with higher levels of VO_{2max} performed better on computerized cognitive tasks assessing sustained attention and working memory and had favourable resting peripheral levels of inflammatory and brain-derived neurotrophic biomarkers. Zhu et al. [88] measured cardiorespiratory fitness (CRF) in 2,747 men and women aged 18 to 30 years and then assessed cognitive function 25 years later. Better verbal memory and faster psychomotor speed at ages 43 to 55 years were clearly associated with better CRF 25 years earlier. One explanation for this is an increased volume of prefrontal and temporal grey matter, as well as anterior white matter, seen in individuals with high levels of aerobic fitness [89]. Alternatively, Etnier et al. [90], looked at Event-related Brain Potential (ERP), a time-locked index of neuroelectric activation that is associated with specific cognitive processes, and observed larger amplitude and shorter latency P3s across a variety of cognitive tasks in individuals with high aerobic fitness compared with unfit individuals. These results collectively indicate that greater amounts of physical activity or aerobic fitness are generally beneficial to cognitive processes. The distinction between aerobic fitness and physical activity is important here and both should be encouraged and incentivised (see Chapter 4).

In conclusion, when determining the aerobic capacity required for effective performance in a given role, the aerobic demand of discrete criterion tasks is just one component, but at this point in time, it is the only component that has been – and can be – reliably assessed. An impact of aerobic capacity on cognition, repeated/extended performance, recovery, etc is apparent, but the level of aerobic capacity at which functioning in these areas is optimised (or even sufficient) is unknown and therefore a bona fide occupational requirement above that measured for a discrete task cannot be established. Further research in this area is required, with an emphasis on establishing not just the relationship between aerobic capacity and multi-dimensional operational effectiveness, but also determination of the minimum level of functioning that is acceptable and or can be justifiably expected/desired. Individual differences and potential protective factors should also be explored.

8.3 TIMELINES/REVISION

8.3.1 Typical Timeline for Development of a Best Practice Physical Employment Standard

Development of an evidence-based, valid and reliable, job-related Physical Employment Standard is not a short or quick process. Most military organisations surveyed for this report took around 4 years to develop and implement their PES. Projects were led by interdisciplinary project teams of scientists, practitioners and subject matter experts, and executed by data collection teams of varying size. Table 8-1 provides an indication of the duration and project team size for various international military PES projects.

Table 8-1: Indications of the Duration and Project Team Size for Military PES Projects from Different Countries.

	Duration	Project Management Team Size	Data Collection Team Size
US Marines	4 years	10	5
NZ Army	4 years	8	2
Norwegian Army	4 years	5	2

	Duration	Project Management Team Size	Data Collection Team Size
US Army	3 years	6	15 – 20
Australian Army	4 years	6-8	6 – 8
Royal Australian Navy	4 years	5	4 – 6
Dutch Armed Forces (pre-screening test)	1 year	3	3
UK Armed Forces Ground Close Combat Roles	3.5 years	10	8
British Army Non-Ground Close Combat Roles	3 years	10	8
Canadian Armed Forces	4 years	15	6
Danish Armed Forces	2 years	10	4

The variation, particularly in data collection team size, reflects in part the respective budgets and resource availability of each nation. Where resource is limited, it is possible that where representative military tasks are common between nations: the results of one nation’s task analysis could potentially be used by another nation, thus saving considerable time and money. This report and the common military tasks presented therein can serve as the basis for such discussion to commence. In the future, a more directed NATO-recommended PES protocol should be developed, which nations who share the same common military tasks could adopt.This would require the agreement and implementation of comparative methods in data collection and data processing and analysis to facilitate the application of established relationships between predictive and task based tests. One developing example is the use of the Isometric Mid-thigh pull to predict performance on various casualty evacuation tasks. Currently the UK, Canada, and Australia are using the same device, and data capture and analysis to establish these relationships and to determine if they can apply across nations.

8.3.2 Best Practice for When and How to Revisit and Revise Physical Employment Standards

In a continuously advancing technological workforce, occupational tasks and demands are evolving and changing. This will require previously developed and implemented physical employment standards to be revisited, and in certain situations revised. The question often proposed to the PES researcher is, how often should the physical employment standard be revisited? However, perhaps the more important question is, in what circumstances should the physical employment standard be revisited?

Any significant change in the physical demands of occupational tasks will require a PES to be revisited or revised. This could be in the form of modification of operational procedures, implementation and deployment of new equipment, or addition of new roles to the job description [91]. A change in environment or conditions, or a safety concern, would also require physical employment standards to be revisited. A new requirement to perform the same task at altitude [92] or in extreme heat [93], thus potentially requiring a higher aerobic capacity, would require a review of the new demand and possible adjustment of the minimal physical capacity required to perform the specific task.

As soon as changes in procedures, training, equipment deployment, or job description are detected, the physical employment standard should be revisited, and in the case where the physical demand of the occupation has been modified, the implemented physical performance standard should be revised [94]. This requires ongoing oversight of job requirements and the continued relevance of a nation's PES. Responsibility for this should be documented at the PES implementation stage. In the absence of regular, ongoing overview, or in situations where no notable specific changes to procedures, equipment or job description have occurred, a review of the relatedness of the physical employment standard to occupational demands should be performed within 5 – 8 years of test validation [95].

When criterion tasks change, such as when new equipment is introduced, large data sets are required in order to revalidate the task to test relationship. An advantage of criterion-referenced standards is that adjustments to the task can be reflected by equivalent adjustments to the assessment. When standards are based on composite scoring systems, a pass score is based on cumulative scores over a number of assessments. This may be considered an effective and balanced method of assessment as it can allow for scenarios in which personnel excel in one component of the assessment but perform below the required level in another component. However, this results in having employees that may not have the required physical capacity for all components of their job, which may impair job performance and place the individual at an elevated risk of injury.

As stated above, modification of a predictive test and/or test standard will be influenced by changes in training, policy and operation requirements, such as job task standards and criteria. A recent example of policy change that has initiated the development/revalidation of PES for many nations is the lifting of the exclusion of women from ground close combat roles. Such a change should not in itself necessitate a change in test or test standard, merely an assurance that the test and standard still reflect the demands of the role, regardless of gender.

Ongoing, transparent oversight of military roles and the associated PES provides the end-user with reassurance that the test(s) and standard(s) remain current and helps ensure that any required changes are received positively rather than with "they got it wrong" or "not another change". The details of any changes made to tests and/or standards, including the originator, purpose, procedures, implementation, and impact (on success rates, manning, etc.), should be captured diligently for auditable purposes.

Currently, no nations have a set timeline or formal process for PES review. The US Army is constantly reviewing the impact of the OPAT through ongoing monitoring of injuries, training failures, attrition and job migration. The British Army recognises that with its imminent introduction of new PES, an enduring process will be required to monitor and respond to any changes in equipment or practices. The Danish Armed Forces plan to evaluate the introduction of their new PES after one year and the Norwegian Army will review their testing as and when required, based on the above guidance. The New Zealand Army will review its Land Combat Fitness Test in 2019, which will be 5 years since its implementation. The Canadian Armed Forces have already begun re-validating their Critical and Essential tasks, through focus groups, as their PES was introduced in 2013. Finally, the Dutch Armed Forces intend to have a working group come together every two years to evaluate if the essential job components that have been operationalized into medical and physical capacity standards for entry into the Dutch Armed Forces are still valid.

Continuous monitoring of occupational demands and test- and cut scores of physical performance standards by the organisation, in cooperation with researchers, would be the optimal way to monitor the effectiveness and applicability of implemented physical employment standards [91]. The advantage of having an in-house R&D team dedicated to PES development and management, as opposed to contracting a university, for example, is that continuous internal review is more feasible. An in-house team will be better placed to track the activities of the military, be it combat or humanitarian missions and operations, and advise when a change in PES is required. Once identified, researchers would ideally be embedded with the military members on operations that deviate from the norm, such as a new environment or the introduction of a new ship or equipment.

In large organisations like the military, employing tens (if not hundreds) of thousands of people, continuous monitoring is a time consuming and costly approach. An alternative, more realistic approach would involve the creation of a PES working group (with consistent, enduring membership) that conducts qualitative annual review, to guide a quantitative review/evaluation as and when required.

Electronic storage of PES results in an interrogative database will allow regular (at least annual) analysis of data, which can also be used to inform the timing of test review and answer questions relating to injury, pass rates, longitudinal trends, etc. The Canadian Armed Force review their PES results every 2 years and adjust the norm-referenced incentives as appropriate.

8.3.3 Best Practice for Ensuring Consistent and Accurate Administration of the Approved PES

Most nations do not follow a formalised process for ensuring that tests are being run in the way they were originally intended and note that deviation from intended methodology is common. Periodic review of how tests are administered is recommended. The US Marines conduct MOS-specific physical standard inspections every 2 years, in a planned fashion, but acknowledge that random spot-checks may be more effective. In the Norwegian Army, Test Leaders must be certified to deliver the required protocols correctly and the Dutch Armed Forces intend to have an expert team regularly visit the institute responsible for administration of their pre-entry screening to ensure quality control.

In Canada, an operations manual was developed and is now undergoing its fifth revision in six years to ensure that every question or concern is addressed. This can be as detailed as calibrating the resistance when dragging on various surfaces, or updates with regard to advances in new technology used to measure and score test results. The Canadian Armed Forces have in-house civilian Physical Training Instructors with an education in Kinesiology attained from a secondary institution, and the Canadian Society of Exercise Physiology have certified their ability to administer tests such as maximal protocols and anthropometrics. In addition, these 250 field staff attend a training centre for one week during induction, to be appropriately trained on delivering the FORCE evaluation (PES), and any military staff who support this fitness testing are also trained by civilians.

8.4 INTERNATIONAL PERSPECTIVE OF THE SOLDIER FIRST CONCEPT

8.4.1 Introduction

The Solider First concept is a topic that often arises at the start of a military PES test development process. It typically refers to the principle that regardless of age, trade or rank, a uniformed military member, be they a chef, musician or infantry, is a "soldier first." Throughout 2013 – 2019 the NATO RTG HFM-269 members generally agreed that a soldier first principle existed in each nation in relation to the development of Physical Employment Standards (PES), but that this was not always well defined.

The term 'Soldier First' can represent two interrelated concepts, firstly it can be used to represent an ethos where personnel are in uniform and therefore part of an organisation larger than their unit or regiment akin to a corporate identity. This is reinforced through common/similar entry tests, basic training, the wearing of uniform and the traditions of their respective service. Service personnel often talk of an Esprit des Corps and whilst some of their tasks/roles may appear similar to those conducted by civilian counterparts, it is being in the Service, be it Navy, Marines, Army or Air Force, that makes the job different.

Secondly and of more direct practical relevance when designing Physical Employment Standards is that most personnel will have to be capable of operating in field conditions and in a state of emergency, and therefore need to be able to carry out common soldiering or common military tasks. The common military tasks are the foundation of a PES, once the physical elements have been identified and quantified. It is not surprising that NATO nations share similarity in their defined common tasks for "solider first". However, what is different, is who these tasks apply to, i.e., who is required to be capable of acting as a "soldier first" (Figure 8-7).

**CASE STUDY:
SOLDIER FIRST PRINCIPLE**

The Universality of Service Principle or "soldier first" principle holds that CAF members (Army/Navy/ or Air Force) are liable to perform general military duties and common defense and security duties, not just the duties of their military occupation or occupational specification. This may include, but is not limited to, the requirement to be physically fit, employable and deployable for general operational duties. Universality of Service has existed since at least 1985. It is imposed by section 33(1) of the National Defence Act, which states that all Regular Force members are "at all times liable to perform any lawful duty." This legislative imperative means that a member who cannot "at all times … perform any lawful duty" may not serve in the Regular Force except during a limited period of recovery from injury or illness. This is to allow for transition back into military service for those whose impairment is only temporary, or out of the military and into civilian life for those permanently unable to meet the liability imposed by the Act. In addition, Section 33(2) of the National Defence Act states that Reserve Force members "may be called out on service to perform any lawful duty other than training at such times and in such manner as any regulations or otherwise are prescribed by the Governor in Council." Since the Primary Reserve is given the role of directly supporting the Regular Force, operational effectiveness requires these Reservists to meet the same Universality of Service standards as their Regular Force counterparts.

Figure 8-7: Case Study: Soldier First Principle.

The definitions of these common soldiering tasks can be explained in more detail by reviewing the origins within each Army (Figure 8-8).

Loaded March: Personnel may operate either as standalone entities, or in support or alongside direct combat forces, but not necessarily in a combatant role. As such all personnel need to be self-sufficient so be able to carry their own equipment be it on a loaded march or simply moving their kit from the airhead to accommodation/workplace, which in the field may be a considerable distance over rough terrain. This load can be in a rucksack, or personal protective equipment such as a fragmentation vest, tactical vest, and communication equipment.

Casualty Evacuation: A key requirement for service personnel is the ability to evacuate casualties. This generally comprises two elements the first being the removal from immediate danger, and the second is movement to a place of evacuation/secondary care. For example dragging a casualty into cover followed by the movement of foot by stretcher to an evacuation point/aid station. These five nations (Figure 8-8) have a variety of scenarios for which their CASEVAC is undertaken which would be the result of SME reports. For the Canadian Armed Forces this involves a Vehicle Extrication from a pick-up truck, a two man casualty carry from a 3rd story building, and a stretcher carry in a domestic aid to civilians. Other nations cites a casualty occurring when dismounted, and the past 10 years of joint combat missions result in the expansion of vehicle borne operations where a casualty situation is likely to occur during transit in a military vehicle. This can be defined as a lateral drag out of the front or back or the vertical extraction from a turret or an overturned vehicle. Alternatively, personnel might be simply required to assist in the movement of casualties around a firm base on a stretcher.

SOLDIER FIRST

Common soldiering tasks of 5 nations

What are the common soldiering tasks				
Australian Army	Canadian Army	New Zealand Army	United States Army	British Army
Loaded march Fire and movement Box lift and place Carry	Loaded March Escape to cover Sandbag fortification Fence construction Stretcher carry Picking and digging Urban Ops Casevac Vehicle Casevac	Loaded march Fire and manoeuvre Lift and place Lift and carry	Tactical foot march Move under fire Build sandbag bunker Casualty drag Vehicle Casevac	Loaded March Fire and Move Water can carry Repeated Lift and carry Casualty Drag Vehicle Casevac
Who must complete these tasks				
Australian Army	Canadian Army	New Zealand Army	United States Army	British Army
Artillery Armoured Engineers Infantry	All Canadian Army	Artillery Armoured Engineers Infantry	Seven combat jobs	Armoured Infantry

Figure 8-8: Common Soldiering Tasks of Five Nations.

Fire and Movement/Move Under fire: The infantry's job is to deliberately close with the enemy, often requiring them to conduct patrolling with load often followed by fire and movement. Personnel supporting the infantry (e.g., fire support controllers, combat medics) will have to patrol with the infantry carrying a similar load moving at the same pace. While the role of these personnel may not require them to close with the enemy, during a patrol they may come under fire requiring them to conduct offensive or defensive fire and movement without endangering their compatriots or their own safety. Personnel in a firm base or camp may also come under fire requiring them to conduct effective fire movement regardless of their role, or seek shelter from fire.

A caveat to this task for some nations' Chaplains, who may accompany patrols as a means of moving between locations, will not carry weapons nor ammunition. However, they still have a requirement to be self-sufficient so as not to be a burden on the patrol, meaning they would still need to keep to the patrol's pace and keep themselves safe if they come under fire (i.e., move to cover).

Manual Handling: Logistics are key to military operations and involve some element of personnel manually lifting, shifting and loading stores (fuel, food, water and ammunition) and movement and erection of equipment (i.e., tents and radio masts). Whilst some specialists' roles may have a high manual handling requirement, practically all roles will need to move loads between 10 – 20 kg or 22 – 44 lbs.

The following sections describe the approaches taken by the Australian Defence Force, the New Zealand Army, the Canadian Armed Forces, the United States Army, and the British Army in developing PES, with consideration to the soldier first concept.

8.4.2 Australian Defence Force

The following information describes the approach taken by the Australian Defence Force in developing a PES designed to represent the demands of military tasks and ensure all personnel have the appropriate physical capacity to safely and effectively carry out essential tasks. These were developed by the Defence Science and Technology Group and is reported in two technical reports [96], [97].

The overall aim of the Australian Army PES research was to develop a suite of core physical capacity based assessments with variable performance standards set according to both employment category and generic military requirements.

Two PES categories, based upon operational roles, were defined and PES designed, reflecting four key physical capacities:

- All Corps Soldier (ACS): The ACS role is based on personnel operating effectively in a defensive environment or firm base, with low-medium threat and action focussed on individual and/or force preservation. The typical load worn by the ACS was defined as fighting order, with a mass of 22.7 kg. Fourteen key criterion tasks were identified for the ACS, with four criterion tasks down-selected and subsequently used to develop PES tests (loaded march, break contact drill, box lift and place, stretcher carry).
- Combat Arms (CA): The CA role is based on operating forward as part of a combined arms team, or in a high threat environment, expecting direct enemy contact and may require offensive action. The identified 'base' loads were fighting order (22.7 kg) and marching order (38.4 kg). Twelve criterion tasks were identified for the CA soldier, with four criterion tasks down-selected and subsequently used to develop PES tests (loaded march, fire and movement, box lift and place, stretcher carry).
- Each employment category was reviewed with reference to the ACS and CA tests and standard for the four physical capacities identified (aerobic, anaerobic, muscular strength, muscular endurance) to ensure adequate cover for role-related criterion tasks. Where the ACS and CA PES did not adequately assess required physical capacity for an employment category, or if any task required a

higher standards and/or the development of an additional tests were developed and standard to adequately reflect the criterion military tasks relevant to that employment category.

8.4.3 New Zealand Army

The following section outlines the approach taken by the New Zealand Army in developing a land combat fitness test to ensure assessment of combat readiness reflects current operational demands and is applicable across all trades. The work is described in a technical report published by the Defence Technology Agency, New Zealand [98].

The Land Combat Fitness Test (LCFT) project was initiated to:

- Develop an operational fitness test that confirms an individual's physical readiness or otherwise to perform the required tasks at Directed Level of Operational Capability (DLOC);
- Encourage ongoing task-specific physical training to meet a relevant and validated level of performance – more soldiers available for deployment;
- Identify areas of weakness to guide future preparation; and
- Ultimately enhance operational performance and reduce injury.

All physical tasks performed by the "All Arms Soldier", irrespective of trade, gender, age or training, were identified via seven focus group discussions involving SME input and doctrinal review to identify key military tasks. These key tasks were then classified and ranked to determine the most physically demanding within each of four physical component categories:

- Aerobic power: sustained whole-body endurance tasks (e.g., patrolling, digging).
- Anaerobic power: short duration, whole-body high intensity tasks (e.g., fire and manoeuvre).
- Muscular endurance: repeated lifting tasks and sustained carrying tasks (e.g., jerry can re-supply, stretcher carry).
- Muscular strength: discrete, heavy weight lifting tasks (e.g., unloading/loading a generator).
- Nine criterion tasks were initially identified and subsequently narrowed down to five tasks and finally four tests, after the casualty drag became redundant following prediction by the successful performance of the Lift and Place test. The report acknowledges the work drew heavily on the Australian PES work, with data on the physical demands of soldiering tasks common to both nations being generously shared by DSTO. In addition, a detailed physical demands analysis was not conducted on every task, as identification of the most physically demanding tasks was straightforward and uncontested, and the task could be the test, saving considerable time and money.

8.4.4 Canadian Armed Forces

The following section outlines the approach taken by Canadian Armed Forces in developing the Common Military Task Fitness Evaluation (CMTFE) as the Bona fide Occupational Requirement for Employment in the Canadian Forces. The work was carried out by the Human Performance Research and Development Team – Canadian Forces Morale and Welfare Services, and the summary below is based on a synopsis of two technical reports [99], [100].

For all Canadian Forces (i.e., a tri-service approach), the essential components of the job were identified by consulting literature from the last 20 years. This identified 72 tasks (e.g., casualty evacuation, securing an area and establishing temporary camp) which were reviewed by two panels of subject matter experts considering both expeditionary and domestic deployments.

All identified tasks were generic military tasks, based on the principle of Universality of Service cited in Canadian employment law [101], often referred to as the “soldier first” principle. This requires that all personnel be able to perform general military duties and common defence and security duties in addition to their military occupation or occupational speciality.

An expert panel subsequently identified 13 tasks as “common and essential” to employment in the Canadian Forces. The physical demands of these 13 common and essential tasks were measured, and a systematic process was employed to reduce the list of 13 tasks to those which are most physically demanding, the theory being that if you can perform those which are most physically demanding, you have the capacity to perform all 13 tasks. These final six, most demanding tasks, would be used to define and develop the CMTFE: Fence construction, Escape to cover, Sandbagging, Picking/digging, Vehicle CASEVAC, Stretcher carry. In addition to being competent on these universal CAF tasks, the Canadian Army is also evaluated on their ability to perform two, 2 man casualty evacuations from a three story building, and a loaded march.

8.4.5 United States Army

The following section outlines the approach taken by the US Army to develop the baseline physical readiness requirements for key physically demanding, commonly occurring, critical Warrior Tasks and Battle Drills (WTBD), and Common Soldier Tasks (CST). A two-pronged approach was taken: one to develop a screening tool to select recruits ready to train for Army Military Occupational Specialties (MOSs) and a second to develop a physical fitness test for operational Army soldiers that was predictive of physically demanding WTBDs and CSTs.

The first effort resulted in the development of the Occupational Physical Assessment Test (OPAT). A job analysis of seven combat jobs identified 32 critical physically demanding tasks. Eight tasks were down-selected as the most demanding, critical and frequently performed tasks across the combat arms MOSs. Five of the eight tasks are CSTs (casualty evacuation from a vehicle hatch, casualty drag, build a sandbag bunker, move under fire and tactical foot march) [96]. A concurrent validation study was conducted on fully trained soldiers and resulted in the four-event OPAT (standing long jump, seated power throw, strength deadlift, and interval aerobic run) [97], [98]. A subsequent longitudinal validation study confirmed the ability of the OPAT to predict physically demanding task performance at the end of initial entry training in new recruits ($R^2 = 0.74$) and was used to help set the test standards [99]. Following additional job analysis, the OPAT was expanded to include all Army occupations. Beginning January 2017, all personnel entering the Army were required to pass the OPAT at a level specific to their assigned occupation (heavy, significant, or moderate physical demand) before they would be allowed to start initial entry training [102], [103].

The second effort was to develop a physical and operational readiness test to improve individual soldier combat readiness and transform the culture of Army fitness. The goal was to ensure all soldiers were physically ready to deploy and fight. Subject Matter Experts (SMEs) identified five basic warrior tasks, including shoot, move, communicate, survive, and adapt. These warrior task categories were expanded into WTBDs. A WTBD proxy simulation test (WTBD-PST) was developed and performed wearing a fighting load. The following task simulations were performed in sequence:

1) Prepare a fighting position;
2) Move over, under, around and through obstacles;
3) React to man-to-man contact; and
4) Casualty extraction and evacuation.

To increase the fidelity of the WTBD-PST, a 1600m road march (the fifth WTBD/CST) was conducted just prior to PST execution to simulate a movement to contact [104], [105].

324 men and 278 women soldiers participated in a validation study to evaluate the ability of 23 physical fitness tests to predict performance on the WTBD-PST. Seven field expedient physical fitness test events (sled drag, 2-mile run, standing power throw, deadlift, 400-m sprint, sled push, and kettlebell squat) were found to be 72% predictive of WTBD-PST performance. Muscle power and aerobic fitness were the strongest predictors of WTBD-PST completion time. A second validation study was conducted with fewer predictors and all predictors were executed with a limited rest between events. Based on these results, a six event Army Combat Fitness Test (ACFT) was selected that was 80% predictive of WTBD-PST performance (Personal communication, Dr W.B. East, US Army Center for Initial Military Training, November 1st, 2018). The ACFT is conducted in order with limited rest between events and includes a three repetition maximum hexbar deadlift, standing power throw, hand release push-up, sprint-drag-carry shuttles, leg tuck and 2 mile run tests. Passing scores are based on the physical demands of the soldier's job and are age and gender-neutral. Extensive pilot work will continue until October 2019 when all soldiers will be required to begin taking the test. Following a one year train-up period, all soldiers must pass the ACFT beginning October 2020 [105].

8.4.6 British Army

The British Army PES for Ground Close Combat (GCC) roles (Royal Armoured Corps and Infantry) came into service in April 2019 and provides an in-service test, the Role Fitness Test (Soldier) [RFT(S)], and an empirically linked point-of-selection test, the Role Fitness Test (Entry) [RFT(E)]. The process of developing PES for the 195 non-GCC PES roles [106] commenced in September 2018. As part of the development, the issue of soldier first tasks was discussed, but to-date no decision has been made.

8.5 REFERENCES

[1] Ministry Of Defence. (2018). UK Armed Forces Biannual Diversity Statistics 1 October 2018. Published 13 December 2018. Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/763676/1_October_2018_Biannual_Diversity_Statistics.pdf.

[2] American College of Obstetricians and Gynecologists. (2018 April). Employment Considerations During Pregnancy and the Postpartum Period. ACOG Committee Opinion, Number 733.

[3] Clapp III, J.F., Kim, H., Burciu, B., Schmidt, S., Petry, K., and Lopez, B. (2002). Continuing Regular Exercise During Pregnancy: Effect of Exercise Volume on Fetoplacental Growth. American Journal of Obstetrics and Gynecology, 186(1), 142-147.

[4] Artal, R., and O'Toole, M. (2003). Guidelines of the American College of Obstetricians and Gynecologists for Exercise During Pregnancy and the Postpartum Period. British Journal of Sports Medicine, 37(1), 6-12.

[5] Clapp III, J.F. (2008). Long-term Outcome after Exercising Throughout Pregnancy: Fitness and Cardiovascular Risk. American Journal of Obstetrics and Gynecology, 199(5), 489-e1.

[6] Tikkanen, J., and Heinonen, O.P. (1991). Maternal Hyperthermia During Pregnancy and Cardiovascular Malformations in the Offspring. European Journal of Epidemiology, 7(6), 628-635.

[7] Kuhrt, K., Harmon, M., Hezelgrave, N.L., Seed, P.T., and Shennan, A.H. (2018). Is Recreational Running Associated with Earlier Delivery and Lower Birth Weight in Women Who Continue to Run During Pregnancy? An International Retrospective Cohort Study of Running Habits of 1293 Female Runners during Pregnancy. BMJ Open Sport and Exercise Medicine, 4(1), e000296.

[8] Dunning, K., Lemasters, G., Levin, L., Bhattacharya, A., Alterman, T. and Lordo, K. (2003). Falls in Workers During Pregnancy: Risk Factors, Job Hazards, and High Risk Occupations. American Journal of Industrial Medicine, 44: 664-672.

[9] Inanir, A., Cakmak, B., Hisim, Y., and Demirturk, F. (2014). Evaluation of Postural Equilibrium and Fall Risk During Pregnancy. Gait and Posture, 39(4), 1122-1125. doi:10.1016/j.gaitpost.2014.01.013.

[10] Butler, E.E., Colon, I., Druzin, M.L., and Rose, J. (2006). Postural Equilibrium During Pregnancy: Decreased Stability with an Increased Reliance on Visual Cues. American Journal of Obstetrics and Gynecology, 195(4), 1104-1108. doi:10.1016/j.ajog.2006.06.015.

[11] Hills, N.F., Graham, R.B., and Mclean, L. (2018). Comparison of Trunk Muscle Function Between Women with and Without Diastasis Recti Abdominis at 1 Year Postpartum. Physical Therapy, 98(10), 891-901.

[12] Branco, M., Santos-Rocha, R., and Vieira, F. (2014). Biomechanics of Gait during Pregnancy. The Scientific World Journal, 2014, 527940. doi. /10.1155/2014/527940.

[13] Foti, T., Davids, J.R., and Bagley, A. (2000). A Biomechanical Analysis of Gait During Pregnancy. Journal of Bone and Joint Surgery, 82(5), 625-632.

[14] Huang, T.H., Lin, S.C., Ho, C.S., Yu, C.Y., and Chou, Y.L. (2002). The Gait Analysis of Pregnant Women. Biomedical Engineering: Applications, Basis and Communications, 14(02), 67-70.

[15] Hagan, L., and Wong, C.K. (2010). Gait in Pregnant Women: Spinal and Lower Extremity Changes from Pre-to Postpartum. Journal of Women's Health Physical Therapy, 34(2), 46-56.

[16] Ravelli, G.P., Stein, Z.A., and Susser, M.W. (1976). Obesity in Young Men after Famine Exposure in Utero and Early Infancy. New England Journal of Medicine, 295(7), 349-353.

[17] Roseboom, T.J., Van Der Meulen, J.H., Ravelli, A.C., Osmond, C., Barker, D.J., and Bleker, O.P. (2001). Effects of Prenatal Exposure to the Dutch Famine on Adult Disease in Later Life: An Overview. Twin Research and Human Genetics, 4(5), 293-298.

[18] Normile, D. (2018). Staying Slim During Pregnancy Carries a Price. Science 2018;361:440.

[19] United States Army Public Health Command. U.S. Army Pregnancy and Postpartum Physical Training Implementation Guide. Directorate of Health Promotion and Wellness, Aberdeen Proving Ground, Maryland. June 2010.

[20] Butte, N.F., Wong, W.W., Treuth, M.S., Ellis, K.J., and O'Brian Smith, E. (2004). Energy Requirements During Pregnancy Based on Total Energy Expenditure and Energy Deposition. The American Journal of Clinical Nutrition, 79(6), 1078-1087.

[21] Gilmore, L.A., Butte, N.F., Ravussin, E., Han, H., Burton, J.H., and Redman, L.M. (2016). Energy Intake and Energy Expenditure for Determining Excess Weight Gain in Pregnant Women. Obstetrics and Gynecology, 127(5), 884.

[22] Wigle, D.T., Arbuckle, T.E., Turner, M.C., Bérubé, A., Yang, Q., Liu, S., and Krewski, D. (2008). Epidemiologic Evidence of Relationships Between Reproductive and Child Health Outcomes and Environmental Chemical Contaminants. Journal of Toxicology and Environmental Health, Part B, 11(5-6), 373-517.

[23] Smith, R., Shakespeare, J., Williams, Z., Knight, M., and Foster, C. (2017). Physical Activity for Pregnant Women: An Infographic for Healthcare Professionals. The British Journal of General Practice: The Journal of the Royal College of General Practitioners, 67(663), 460-460.

[24] Fried, M., and Duffy, P.E. (1998). Maternal Malaria and Parasite Adhesion. Journal of Molecular Medicine, 76(3-4), 162-171.

[25] Khan, S., Woodruff, E.M., Trapecar, M., Fontaine, K.A., Ezaki, A., Borbet, T.C., ... and Sanjabi, S. (2016). Dampened Antiviral Immunity to Intravaginal Exposure to RNA Viral Pathogens Allows Enhanced Viral Replication. Journal of Experimental Medicine, 213(13), 2913-2929.

[26] United States Department of the Army. Army Regulation 40-501. Standards of Medical Fitness. 14 June 2017.

[27] Casey, E.K., and Temme, K. (2017). Pelvic Floor Muscle Function and Urinary Incontinence in the Female Athlete. The Physician and Sports Medicine, 45(4), 399-407.

[28] Bø, K., Artal, R., Barakat, R., Brown, W., Davies, G.A., Dooley, M., ... and Kinnunen, T.I. (2016). Exercise and Pregnancy in Recreational and Elite Athletes: 2016 Evidence Summary from the IOC Expert Group Meeting, Lausanne. Part 1 – Exercise in Women Planning Pregnancy and Those Who are Pregnant. British Journal of Sports Medicine, 50(10), 571-589.

[29] Noon, M.L., and Hoch, A.Z. (2012). Challenges of the Pregnant Athlete and Low Back Pain. Current Sports Medicine Reports, 11(1), 43-48.

[30] Vermani, E., Mittal, R., and Weeks, A. (2010). Pelvic Girdle Pain and Low Back Pain in Pregnancy: A Review. Pain Practice, 10(1), 60-71.

[31] Strowbridge, N.F. (2005). Gender Differences in the Cause of Low Back Pain in British Soldiers. Journal of the Royal Army Medical Corps, 151(2), 69-72.

[32] Sperstad, J.B., Tennfjord, M.K., Hilde, G., Ellström-Engh, M., and Bø, K. (2016). Diastasis Recti Abdominis During Pregnancy and 12 Months after Childbirth: Prevalence, Risk Factors and Report of Lumbopelvic pain. British Journal of Sports Medicine, 50(17), 1092-1096.

[33] Liaw, L.J., Hsu, M.J., Liao, C.F., Liu, M.F., and Hsu, A.T. (2011). The Relationships between Inter-Recti Distance Measured by Ultrasound Imaging and Abdominal Muscle Function in Postpartum Women: A 6-month Follow-up Study. Journal of Orthopaedic and Sports Physical Therapy, 41(6), 435-443.

[34] Salari, P., and Abdollahi, M. (2014). The Influence of Pregnancy and Lactation on Maternal Bone Health: A Systematic Review. Journal of Family and Reproductive Health, 8(4), 135.

[35] Kovacs, C.S. (2016). Maternal Mineral and Bone Metabolism During Pregnancy, Lactation, and Post-Weaning Recovery. Physiological Reviews, 96(2), 449-547.

[36] Affinito, P.I., Tommaselli, G.A., Di Carlo, C.O., Guida F.R., Nappi, C.A. (1996). Changes in Bone Mineral Density and Calcium Metabolism in Breastfeeding Women: A One Year Follow-Up Study. The Journal of Clinical Endocrinology and Metabolism, 81(6), 2314-2318.

[37] Little, K.D. (1998). Self-selected Recreational Exercise has no Impact on Early Postpartum Lactation-Induced Bone Loss. Medicine and Science in Sports Exercise, 30(6), 831-836.

[38] Lovelady, C.A., Bopp, M.J., Colleran, H.L., Mackie, H.K., and Wideman, L. (2009). Effect of Exercise Training on Loss of Bone Mineral Density During Lactation. Medicine and Science in Sports Exercise, 41(10), 1902-7.

[39] Woolhouse, H., Perlen, S., Gartland, D., and Brown, S.J. (2012). Physical Health and Recovery in the First 18 Months Postpartum: Does Cesarean Section Reduce Long-Term Morbidity? Birth, 39(3), 221-229.

[40] Brown, S.J., Gartland, D., Donath, S., and Macarthur, C. (2012). Fecal Incontinence During the First 12 Months Postpartum: Complex Causal Pathways and Implications for clinical practice. Obstetrics and Gynecology, 119(2), 240-249.

[41] Boyle, R., Hay-Smith, E.J.C., Cody, J.D., and Mørkved, S. (2014). Pelvic Floor Muscle Training for Prevention and Treatment of Urinary and Fecal Incontinence in Antenatal and Postnatal Women: A Short Version Cochrane Review. Neurourology and Neurodynamics, 33(3), 269-276.

[42] Sherman, R.A., Davis, G.D., and Wong, M.F. (1997). Behavioral Treatment of Exercise-Induced Urinary Incontinence Among Female Soldiers. Military Medicine, 162(10), 690-694.

[43] Hilde, G., Stær-Jensen, J., Siafarikas, F., Engh, M.E., Bø, K. (2013). Postpartum Pelvic Floor Muscle Training and Urinary Incontinence: A Randomized Controlled Trial. Obstetrics and Gynecology. 2013 Dec 1;122(6):1231-8.

[44] Woolhouse, H., Gartland, D., Perlen, S., Donath, S., and Brown, S.J. (2014). Physical Health After Childbirth and Maternal Depression in the first 12 months Post-Partum: Results of an Australian Nulliparous Pregnancy Cohort Study. Midwifery, 30(3), 378-384.

[45] Lee, K.A., Zaffke, M.E., and McEnany, G. (2000). Parity and Sleep Patterns During and After Pregnancy. Obstetrics and Gynecology, 95(1), 14-18.

[46] Demeis, D.K., Hock, E., and McBride, S.L. (1986). The Balance of Employment and Motherhood: Longitudinal Study of Mothers' Feelings About Separation from their First-Born Infants. Developmental Psychology, 22(5), 627.

[47] Mottola, M.F. (2002). Exercise in the Postpartum Period: Practical Applications. Current Sports Medicine Reports, 1(6), 362-368.

[48] Taylor, C.B., Sallis, J.F., and Needle, R. (1985). The Relation of Physical Activity and Exercise to Mental Health. Public Health Reports, 100(2), 195.

[49] Van Der Pligt, P., Willcox, J., Hesketh, K.D., Ball, K., Wilkinson, S., Crawford, D., and Campbell, K. (2013). Systematic Review of Lifestyle Interventions to Limit Postpartum Weight Retention: Implications for Future Opportunities to Prevent Maternal Overweight and obesity Following Childbirth. Obesity Reviews, 14(10), 792-805.

[50] Nascimento, S.L., Pudwell, J., Surita, F.G., Adamo, K.B., and Smith, G.N. (2014). The Effect of Physical Exercise Strategies on Weight Loss in Postpartum Women: A Systematic Review and Meta-Analysis. International Journal of Obesity, 38(5), 626.

[51] Treuth, M.S., Butte, N.F., and Puyau, M. (2005). Pregnancy-Related Changes in Physical Activity, Fitness, and Strength. Medicine and Science in Sports and Exercise, 37(5), 832-837.

[52] Weina, S.U. (2006 June). Effects of Pregnancy on the Army Physical Fitness Test. Military Medicine. 2006 Jun 1;171(6):534-7.

[53] Costello, R.B. (1998). Nutrient Requirements, Body Composition, and Health of Military Women. National Academy of Sciences-National Research Council Washington DC.

[54] Evenson, K.R., Mottola, M.F., Owe, K.M., Rousham, E.K., and Brown, W.J. (2014). Summary of International Guidelines for Physical Activity Following Pregnancy. Obstetrical and Gynecological Survey, 69(7), 407.

[55] McCreight, S.J. (2014). Program Outcomes and Soldier Perceptions of the Army Pregnancy and Postpartum Physical Training (P3T) Program. Uniformed Services University of the Health Sciences Bethesda MD, Bethesda, United States.

[56] Health and Safety Executive. Management of Health and Safety at Work Regulation 1999 (MHSW). Implementation of the Health and Safety Requirements of the Pregnant Workers Directive (92/85/EEC) into GB law. Aa\vailable at: http://www.hse.gov.uk/mothers/law.htm.

[57] Ministry of Defence. JSP 760. Tri-Service Regulations for Leave and Other Types of Absence. Ch. 24. Pregnancy and the Armed Forces Occupational Maternity Scheme (AFOMS). Part 1: Directive. Version 37.

[58] Evans, M.A., and Rosen, L. (1997). Pregnancy Planning and the Impact on Work Climate, Psychological Well-Being, and Work Effort in the Military. Journal of Occupational Health Psychology, 2(4), 353.

[59] Davies, G.A., Wolfe, L.A., Mottola, M.F., and Mackinnon, C. (2003). Joint SOGC/CSEP Clinical Practice Guideline: Exercise in Pregnancy and the Postpartum Period. Canadian Journal of Applied Physiology, 28(3), 329-341.

[60] Cowan, D.N., Jones, B.H., and Shaffer, R.A. (2003). Musculoskeletal Injuries in the Military Training Environment. Military Preventive Medicine: Mobilization and Deployment, 1.

[61] Jones, B.H., Bovee, M.W., Harris III, J.M., and Cowan, D.N. (1993). Intrinsic Risk Factors for Exercise-Related Injuries Among Male and Female Army Trainees. The American Journal of Sports Medicine, 21(5), 705-710.

[62] Anderson, M.K., Grier, T., Dada, E.O., Canham-Chervak, M., and Jones, B.H. (2017). The Role of Gender and Physical Performance on Injuries: An Army Study. American Journal of Preventive Medicine, 52(5), e131-e138.

[63] King, J., Roberts, C., Hard, S., and Ardern, C.L. (2019). Want to Improve Return to Sport Outcomes Following Injury? Empower, Engage, Provide Feedback and Be Transparent: 4 Habits! British Journal of Sports Medicine, May 53(9), 526-527.

[64] Thacker, S.B., Branche, C., Gilchrist, J., Jones, B.H., Sleet, D.A., and Kimsey, C.D. (2000). Exercise-related Injuries Among Women: Strategies for Prevention from Civilian and Military Studies. Morbidity and Mortality Weekly Report: Recommendations and Reports, 13-33.

[65] Ardern, C.L., Glasgow, P., Schneiders, A., Witvrouw, E., Clarsen, B., Cools, A., ... and Mutch, S.A. (2016). 2016 Consensus Statement on Return to Sport from the First World Congress in Sports Physical Therapy, Bern. British Journal of Sports Medicine, 50(14), 853-864.

[66] Stinner, D.J., Burns, T.C., Kirk, K.L., and Ficke, J.R. (2010). Return to Duty Rate of Amputee Soldiers in the Current Conflicts in Afghanistan and Iraq. Journal of Trauma and Acute Care Surgery, 68(6), 1476-1479.

[67] Potter, B.K., and Scoville, C.R. (2006). Amputation is not Isolated: An Overview of the US Army Amputee Patient Care Program and Associated Amputee Injuries. JAAOS-Journal of the American Academy of Orthopaedic Surgeons, 14(10), S188-S190.

[68] Gagnon, P., Spivock, M., Reilly, T., Mattie, P., and Stockbrugger, B. (2015). The FORCE Fitness Profile – Adding a Measure of Health-Related Fitness to the Canadian Armed Forces Operational Fitness Evaluation. The Journal of Strength and Conditioning Research, 29, S192-S198.

[69] Reilly, T., Walsh, E., and Stockbrugger, B. (2019). Reliability of FORCE COMBAT™: A Canadian Army Fitness Objective. Journal of Science and Medicine in Sport, 22(5), 591-595.

[70] Honey, J.M., Godsell, P.A., Peralta-Huertas, J.F., Besemann, M.H., and Spivock, M.D. (2014). User Evaluation Trial of Lower Limb Prosthetic Devices. Surgeon General Report SGR-2013-013.

[71] Besemann, M., Gagnon, P., Godsell, P., Bérubé, C. and Spivock, M. (2013). Six Common Military Task Familiarization in Canadian Armed Forces Lower Extremity Amputees. Canadian Forces Military Veterans Health Research Forum.

[72] Wolski, L. and Wenger, H. (2003). A Guide to Fitness During and After Pregnancy in the CF. Canadian Forces Morale and Welfare Internal Report. https://www.cfmws.com/en/AboutUs/PSP/DFIT/Documents/Publications%20and%20Reports/Fitness%20During%20and%20After%20Pregnancy.pdf.

[73] Ardern, C.L., Glasgow, P., Schneiders, A., Witvrouw, E., Clarsen, B., Cools, A., ... and Mutch, S.A. (2016). 2016 Consensus Statement on Return to Sport from the First World Congress in Sports Physical Therapy, Bern. British Journal of Sports Medicine, 50(14), 853-864.

[74] Blacker, S.D., Wilkinson, D.M., Bilzon, J.L., and Rayson, M.P. (2008). Risk Factors for Training Injuries Among British Army Recruits. Military Medicine, 173(3), 278-286.

[75] Knapik, J., Ang, P., Reynolds, K., and Jones, B. (1993). Physical Fitness, Age, and Injury Incidence in Infantry Soldiers. Army Research Inst of Environmental Medicine, Natick, MA, US.

[76] Knapik, J.J., Darakjy, S., Hauret, K.G., Canada, S., Scott, S., Rieger, W., ... and Jones, B.H. (2006). Increasing the Physical Fitness of Low-Fit Recruits Before Basic Combat Training: An Evaluation of Fitness, Injuries, and Training Outcomes. Military Medicine, 171(1), 45-54.

[77] Knapik, J.J., Sharp, M.A., Canham-Chervak, M., Hauret, K., Patton, J.F., and Jones, B.H. (2001). Risk Factors for Training-Related Injuries Among Men And Women in Basic Combat Training. Medicine and Science in Sports and Exercise, 33(6), 946-954.

[78] Roy, T.C., Knapik, J.J., Ritland, B.M., Murphy, N., and Sharp, M.A. (2012). Risk Factors for Musculoskeletal Injuries for Soldiers Deployed to Afghanistan. Aviation, Space, and Environmental Medicine, 83(11), 1060-1066.

[79] Knapik, J.J., Darakjy, S., Jones, S.B., Marin, R.E., Hoedebecke, E.L., Mitchener, T.A., ... and Jones, B.H. (2007). Injuries and Physical Fitness Before and After a Deployment by the 10th Mountain Division to Afghanistan for Operation Enduring Freedom. Aberdeen Proving Ground MD: US Army Center for Health Promotion and Preventive Medicine.

[80] Knapik, J.J., Spiess, A., Grier, T., Sharp, M.A., Lester, M.E., Marin, R., and Jones, B.H. (2012). Injuries Before and After Deployments to Afghanistan and Iraq. Public Health, 126(6), 498-506.

[81] McGawley, K., and Bishop, D. J. (2015). Oxygen Uptake During Repeated-Sprint Exercise. Journal of Science and Medicine in Sport, 18(2), 214-218.

[82] Johnston, R.D., Gabbett, T.J., Jenkins, D.G., and Hulin, B.T. (2015). Influence of Physical Qualities on Post-Match Fatigue in Rugby League Players. Journal of Science and Medicine in Sport, 18(2), 209-213.

[83] Jones, A.M., and Carter, H. (2000). The Effect of Endurance Training on Parameters of Aerobic Fitness. Sports Medicine, 29(6), 373-386.

[84] Candau, R., Belli, A., Millet, G.Y., Georges, D., Barbier, B., and Rouillon, J.D. (1998). Energy Cost and Running Mechanics During a Treadmill Run to Voluntary Exhaustion in Humans. European Journal of Applied Physiology and Occupational Physiology, 77(6), 479-485.

[85] Derrick, T.R., Dereu, D., and Mclean, S.P. (2002). Impacts and Kinematic Adjustments During an Exhaustive Run. Medicine and Science in Sports and Exercise, 34(6), 998-1002.

[86] Mizrahi, J., Verbitsky, O., Isakov, E., and Daily, D. (2000). Effect of Fatigue on Leg Kinematics and Impact Acceleration in Long Distance Running. Human Movement Science, 19(2), 139-151.

[87] Hwang, J., Castelli, D.M., and Gonzalez-Lima, F. (2017). The Positive Cognitive Impact of Aerobic Fitness is Associated with Peripheral Inflammatory and Brain-Derived Neurotrophic Biomarkers in Young Adults. Physiology and Behavior, 179, 75-89.

[88] Zhu, N., Jacobs, D.R., Schreiner, P.J., Yaffe, K., Bryan, N., Launer, L.J., ... and Bouchard, C. (2014). Cardiorespiratory Fitness and Cognitive Function in Middle Age: The CARDIA Study. Neurology, 82(15), 1339-1346.

[89] Colcombe, S.J., Kramer, A.F., Erickson, K.I., Scalf, P., Mcauley, E., Cohen, N.J., ... and Elavsky, S. (2004). Cardiovascular Fitness, Cortical Plasticity, and Aging. Proceedings of the National Academy of Sciences, 101(9), 3316-3321.

[90] Etnier, J.L., Nowell, P.M., Landers, D.M., and Sibley, B.A. (2006). A Meta-Regression to Examine the Relationship Between Aerobic Fitness and Cognitive Performance. Brain Research Reviews, 52(1), 119-130.

[91] Petersen, S.R., Anderson, G.S., Tipton, M.J., Docherty, D., Graham, T.E., Sharkey, B.J., and Taylor, N.A. (2016). Towards Best Practice in Physical and Physiological Employment Standards. Applied Physiology, Nutrition, and Metabolism, 41(6), S47-S62.

[92] Chapman, R.F. (2013). The Individual Response to Training and Competition at Altitude. Br J Sports Med, 47(Suppl 1), i40-i44.

[93] Cheung, S.S., Lee, J.K.W. and Oksa, J. (2016). Thermal Stress, Human Performance, and Physical Employment Standards. Applied Physiology, Nutrition, and Metabolism, 41:S148-S164, https://doi.org/10.1139/apnm-2015-0518.

[94] Reilly, T.J., Gebhardt, D.L., Billing, D.C., Greeves, J.P., and Sharp, M.A. (2015). Development and Implementation of Evidence-based Physical Employment Standards: Key Challenges in the Military Context. The Journal of Strength and Conditioning Research, 29, S28-S33.

[95] Gebhardt, D.L., and Todd A.B. (2010). Physical Performance Tests. Handbook of Employee Selection. Handbook of Employee Selection, Ed. James L. Farr, Nancy T. Tippins. Taylor and Francis, 277-298.

[96] Doyle, T., Billing, D., Drain, J., Carr, A., Ham, D., Fogarty, A., ... and Savage, R. (2011). Physical Employment Standards for Australian Defence Force Employment Categories Currently Restricted to Women-Part A: Physically Demanding Trade Tasks. Defence Science and Technology Group.

[97] Doyle, T., Billing, D., Drain., J., et al., (2012). Physical Employment Standards for Australian Defence Force Employment Categories Currently Restricted to Women – Part B: Physical Employment Assessments and Standards. DSTO-CR-2011-0523, Defence Science and Technology Organisation, Australia

[98] Kilding, H. and Fordy, G. (2014). Development of a Land Combat Fitness Test for the New Zealand Army. DTA Report 388. Defence Technology Agency, Auckland, New Zealand

[99] Spivock, M., Reilly, T., Newton, P., Blacklock, R., Jaenen, S., (2011). Project FORCE Phase I Report: Identification of Common Essential, Physically Demanding Tasks in the CF. Department of National Defence, Assistant Deputy Minister (Science and Technology), Ottawa.

[100] Reilly, T., Blacklock, R., Newton, P., Olinek, S., O'hearn, K., Spivock, M., (2013). Project FORCE Phase II Report: Physical Demands of Common, Essential, Physically Demanding Tasks in the CF. Department of National Defence, Assistant Deputy Minister (Science and Technology). Ottawa.

[101] Canadian Human Rights Act (R.S.C., 1985, c. H-6), Act current to 2019-05-07 and last amended on 2017-06-19. https://laws-lois.justice.gc.ca/eng/acts/h-6/fulltext.html.

[102] Sharp, M.A., Foulis, S.A., Redmond, J.E., Canino, M.C., Cohen, B.S., Hauret, K., Pacheco, B.M., Gebhardt, D.L., Frykman, P.N. and Zambraski, E. J. (2018). Longitudinal Validation of the Occupational Physical Assessement Test (OPAT) U.S. Army Research Institute of Environmental Medicine, Natick, MA, US.

[103] Foulis, S. A., Redmond, J.E., Warr, B.J., Zambraski, E.J., Frykman, P.N. and Sharp, M.A. (2015). Development of the Occupational Physical Assessment Test (OPAT) for Combat Arms Soldiers. In M.P. Division, (Ed.). Army Research Institute of Environmental Medicine Natick, MA: US, USARIEM TECHNICAL REPORT T16-2: https://dod.defense.gov/Portals/1/Documents/wisr-studies/Army%20-%20MEDCOM%20USARIEM%20Task%20Assessment3.pdf.

[104] Jones, B., Nindl, B., Hauret, K., Degroot, D.W., Grier, T., Hauschild, V., Clearfield, E., Anderson, M., Bushman, T., and Rossi, S. (2015). Development of a New Army Standardized Physical Readiness Test January 201 through December 2013, U.S. Army Public Health Center, Aberdeen Proving Ground, MD, US.

[105] Pierce, J.R., Hauret, K.G., Alemany, J.A., Grier, T.L., Sharp, M.A., Redmond, J.E., ... and Jones, B. H. (2018). Physical Performance on the Occupational Physical Assessment Test (OPAT), Army Physical Fitness Test (APFT), and Relationship to Body Mass Index During Initial Entry Training-OPAT Phase 1 (No. S. 0047229-18a). US Army Public Health Center (APHC), Injury Prevention Division Aberdeen Proving Ground, US.

[106] Myers, S.D., Rayson, M.P., Reed, P.R., Nevola, V.R. and Blacker, S.D. (2017). TIN 3.223 Women in Ground Close Combat (WGCC) All Arms Physical Employment Standards (PES) Phase 1 – Final Report, Defence Human Capability Science and Technology Centre.

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's References	3. Further Reference	4. Security Classification of Document
	STO-TR-HFM-269 AC/323(HFM-269)TP/936	ISBN 978-92-837-2261-8	PUBLIC RELEASE

5. Originator	
	Science and Technology Organization North Atlantic Treaty Organization BP 25, F-92201 Neuilly-sur-Seine Cedex, France

6. Title

Combat Integration: Implications for Physical Employment Standards

7. Presented at/Sponsored by

Final report of HFM-269.

8. Author(s)/Editor(s)	9. Date
Multiple	December 2019

10. Author's/Editor's Address	11. Pages
Multiple	242

12. Distribution Statement

There are no restrictions on the distribution of this document. Information about the availability of this and other STO unclassified publications is given on the back cover.

13. Keywords/Descriptors

Fatigue	Physical performance
Females	Physical employment standards
Fitness	Pregnancy
Gender	Post-partum
Incentivization	Sex
Musculoskeletal injury	Women
Occupational demands	

14. Abstract

NATO RTG HFM-269 included ten nations who had either already lifted, or were planning to lift, exclusions of women joining combat roles within their Armed Forces. A fundamental element that has supported the opening of combat roles to women is the development of role-related, age- and sex- free Physical Employment Standards (PES), which are physical fitness tests derived from the job-role an individual performs. When the physical requirements of a job-role are the same for all personnel, such as a military occupation, then the PES related to these roles will be the same irrespective of a person's sex, race or age. This report contains advice and information on current PES practices across the NATO nations represented, a NATO guide for PES development, consideration for incentivization, biological limits to physical performance, musculoskeletal injury prevention, prolonged military work, and return-to-duty (post-partum women, and post-injury recovery). The NATO RTG HFM-269 report will inform and enhance NATO's military readiness by improving the understanding of how PES is an important tool to support combat integration and more broadly improve health readiness habits, reduce injury and optimise physical performance of both men and women across a diverse range of military settings.

NORTH ATLANTIC TREATY ORGANIZATION

SCIENCE AND TECHNOLOGY ORGANIZATION

BP 25
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Télécopie 0(1)55.61.22.99 • E-mail mailbox@cso.nato.int

DIFFUSION DES PUBLICATIONS
STO NON CLASSIFIEES

Les publications de l'AGARD, de la RTO et de la STO peuvent parfois être obtenues auprès des centres nationaux de distribution indiqués ci-dessous. Si vous souhaitez recevoir toutes les publications de la STO, ou simplement celles qui concernent certains Panels, vous pouvez demander d'être inclus soit à titre personnel, soit au nom de votre organisation, sur la liste d'envoi.

Les publications de la STO, de la RTO et de l'AGARD sont également en vente auprès des agences de vente indiquées ci-dessous.

Les demandes de documents STO, RTO ou AGARD doivent comporter la dénomination « STO », « RTO » ou « AGARD » selon le cas, suivi du numéro de série. Des informations analogues, telles que le titre est la date de publication sont souhaitables.

Si vous souhaitez recevoir une notification électronique de la disponibilité des rapports de la STO au fur et à mesure de leur publication, vous pouvez consulter notre site Web (http://www.sto.nato.int/) et vous abonner à ce service.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE
Streitkräfteamt / Abteilung III
Fachinformationszentrum der Bundeswehr (FIZBw)
Gorch-Fock-Straße 7, D-53229 Bonn

BELGIQUE
Royal High Institute for Defence – KHID/IRSD/RHID
Management of Scientific & Technological Research for Defence, National STO Coordinator
Royal Military Academy – Campus Renaissance
Renaissancelaan 30, 1000 Bruxelles

BULGARIE
Ministry of Defence
Defence Institute "Prof. Tsvetan Lazarov"
"Tsvetan Lazarov" bul no.2
1592 Sofia

CANADA
DGSlST 2
Recherche et développement pour la défense Canada
60 Moodie Drive (7N-1-F20)
Ottawa, Ontario K1A 0K2

DANEMARK
Danish Acquisition and Logistics Organization (DALO)
Lautrupbjerg 1-5
2750 Ballerup

ESPAGNE
Área de Cooperación Internacional en I+D
SDGPLATIN (DGAM)
C/ Arturo Soria 289
28033 Madrid

ESTONIE
Estonian National Defence College
Centre for Applied Research
Riia str 12
Tartu 51013

ETATS-UNIS
Defense Technical Information Center
8725 John J. Kingman Road
Fort Belvoir, VA 22060-6218

FRANCE
O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc
BP 72
92322 Châtillon Cedex

GRECE (Correspondant)
Defence Industry & Research General Directorate, Research Directorate
Fakinos Base Camp, S.T.G. 1020
Holargos, Athens

HONGRIE
Hungarian Ministry of Defence
Development and Logistics Agency
P.O.B. 25
H-1885 Budapest

ITALIE
Ten Col Renato NARO
Capo servizio Gestione della Conoscenza
F. Baracca Military Airport "Comparto A"
Via di Centocelle, 301
00175, Rome

LUXEMBOURG
Voir Belgique

NORVEGE
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
NO-2007 Kjeller

PAYS-BAS
Royal Netherlands Military Academy Library
P.O. Box 90.002
4800 PA Breda

POLOGNE
Centralna Biblioteka Wojskowa
ul. Ostrobramska 109
04-041 Warszawa

PORTUGAL
Estado Maior da Força Aérea
SDFA – Centro de Documentação
Alfragide
P-2720 Amadora

REPUBLIQUE TCHEQUE
Vojenský technický ústav s.p.
CZ Distribution Information Centre
Mladoboleslavská 944
PO Box 18
197 06 Praha 9

ROUMANIE
Romanian National Distribution Centre
Armaments Department
9-11, Drumul Taberei Street
Sector 6
061353 Bucharest

ROYAUME-UNI
Dstl Records Centre
Rm G02, ISAT F, Building 5
Dstl Porton Down
Salisbury SP4 0JQ

SLOVAQUIE
Akadémia ozbrojených síl gen. M.R. Štefánika, Distribučné a informačné stredisko STO
Demänová 393
031 01 Liptovský Mikuláš 1

SLOVENIE
Ministry of Defence
Central Registry for EU & NATO
Vojkova 55
1000 Ljubljana

TURQUIE
Milli Savunma Bakanlığı (MSB)
ARGE ve Teknoloji Dairesi Başkanlığı
06650 Bakanliklar – Ankara

AGENCES DE VENTE

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
ROYAUME-UNI

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council Acquisitions
Montreal Road, Building M-55
Ottawa, Ontario K1A 0S2
CANADA

Les demandes de documents STO, RTO ou AGARD doivent comporter la dénomination « STO », « RTO » ou « AGARD » selon le cas, suivie du numéro de série (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Des références bibliographiques complètes ainsi que des résumés des publications STO, RTO et AGARD figurent dans le « NTIS Publications Database » (http://www.ntis.gov).

NORTH ATLANTIC TREATY ORGANIZATION

BP 25
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Télécopie 0(1)55.61.22.99 • E-mail mailbox@cso.nato.int

SCIENCE AND TECHNOLOGY ORGANIZATION

DISTRIBUTION OF UNCLASSIFIED STO PUBLICATIONS

AGARD, RTO & STO publications are sometimes available from the National Distribution Centres listed below. If you wish to receive all STO reports, or just those relating to one or more specific STO Panels, they may be willing to include you (or your Organisation) in their distribution.

STO, RTO and AGARD reports may also be purchased from the Sales Agencies listed below.

Requests for STO, RTO or AGARD documents should include the word 'STO', 'RTO' or 'AGARD', as appropriate, followed by the serial number. Collateral information such as title and publication date is desirable.

If you wish to receive electronic notification of STO reports as they are published, please visit our website (http://www.sto.nato.int/) from where you can register for this service.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Royal High Institute for Defence – KHID/IRSD/RHID
Management of Scientific & Technological Research for Defence, National STO Coordinator
Royal Military Academy – Campus Renaissance
Renaissancelaan 30
1000 Brussels

BULGARIA
Ministry of Defence
Defence Institute "Prof. Tsvetan Lazarov"
"Tsvetan Lazarov" bul no.2
1592 Sofia

CANADA
DSTKIM 2
Defence Research and Development Canada
60 Moodie Drive (7N-1-F20)
Ottawa, Ontario K1A 0K2

CZECH REPUBLIC
Vojenský technický ústav s.p.
CZ Distribution Information Centre
Mladoboleslavská 944
PO Box 18
197 06 Praha 9

DENMARK
Danish Acquisition and Logistics Organization (DALO)
Lautrupbjerg 1-5
2750 Ballerup

ESTONIA
Estonian National Defence College
Centre for Applied Research
Riia str 12
Tartu 51013

FRANCE
O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc – BP 72
92322 Châtillon Cedex

GERMANY
Streitkräfteamt / Abteilung III
Fachinformationszentrum der Bundeswehr (FIZBw)
Gorch-Fock-Straße 7
D-53229 Bonn

GREECE (Point of Contact)
Defence Industry & Research General Directorate, Research Directorate
Fakinos Base Camp, S.T.G. 1020
Holargos, Athens

HUNGARY
Hungarian Ministry of Defence
Development and Logistics Agency
P.O.B. 25
H-1885 Budapest

ITALY
Ten Col Renato NARO
Capo servizio Gestione della Conoscenza
F. Baracca Military Airport "Comparto A"
Via di Centocelle, 301
00175, Rome

LUXEMBOURG
See Belgium

NETHERLANDS
Royal Netherlands Military Academy Library
P.O. Box 90.002
4800 PA Breda

NORWAY
Norwegian Defence Research Establishment, Attn: Biblioteket
P.O. Box 25
NO-2007 Kjeller

POLAND
Centralna Biblioteka Wojskowa
ul. Ostrobramska 109
04-041 Warszawa

PORTUGAL
Estado Maior da Força Aérea
SDFA – Centro de Documentação
Alfragide
P-2720 Amadora

ROMANIA
Romanian National Distribution Centre
Armaments Department
9-11, Drumul Taberei Street
Sector 6
061353 Bucharest

SLOVAKIA
Akadémia ozbrojených síl gen M.R. Štefánika, Distribučné a informačné stredisko STO
Demänová 393
031 01 Liptovský Mikuláš 1

SLOVENIA
Ministry of Defence
Central Registry for EU & NATO
Vojkova 55
1000 Ljubljana

SPAIN
Área de Cooperación Internacional en I+D
SDGPLATIN (DGAM)
C/ Arturo Soria 289
28033 Madrid

TURKEY
Milli Savunma Bakanlığı (MSB)
ARGE ve Teknoloji Dairesi Başkanlığı
06650 Bakanliklar – Ankara

UNITED KINGDOM
Dstl Records Centre
Rm G02, ISAT F, Building 5
Dstl Porton Down, Salisbury SP4 0JQ

UNITED STATES
Defense Technical Information Center
8725 John J. Kingman Road
Fort Belvoir, VA 22060-6218

SALES AGENCIES

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
UNITED KINGDOM

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council Acquisitions
Montreal Road, Building M-55
Ottawa, Ontario K1A 0S2
CANADA

Requests for STO, RTO or AGARD documents should include the word 'STO', 'RTO' or 'AGARD', as appropriate, followed by the serial number (for example AGARD-AG-315). Collateral information such as title and publication date is desirable. Full bibliographical references and abstracts of STO, RTO and AGARD publications are given in "NTIS Publications Database" (http://www.ntis.gov).

ISBN 978-92-837-2261-8